1983

The molecular characterization of the short interspersed repeated DNA sequences in the bovine genome

Katherine Anne Koren Richardson
Iowa State University

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106

Richardson, Katherine Anne Koren

THE MOLECULAR CHARACTERIZATION OF THE SHORT INTERSPERSED REPEATED DNA SEQUENCES IN THE BOVINE GENOME

Iowa State University PH.D. 1983

The molecular characterization of the short interspersed repeated DNA sequences in the bovine genome

by

Katherine Anne Koren Richardson

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Zoology
Interdepartmental major: Molecular, Cellular, Developmental Biology

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Interdepartmental Major

Signature was redacted for privacy.

For the Graduate College

Members of the Committee:

Signature was redacted for privacy.

Iowa State University
Ames, Iowa

1983

TABLE OF CONTENTS

DEDICATION

This dissertation is dedicated to:

my Father, John Harald Koren

Who instilled my desire for education, formulated my personality, listened, and wrote two letters from his heart. I'm sure there will always be stones to pick up and weeds to pull, but he has taught me how to overcome such obstacles.

my "Pretty Mommy", Giaconda Aida Monti Koren

Whose love, understanding of human nature, kindness to others, and stamina I hope I've inherited.

my husband, Frank Charles Richardson

His philosophy of life has kept me sane and without him this would not have been possible.

my Noni, Ida Monti

She has been my roommate, friend, and grandmother. Everyone should have a person like my Noni - she's special!

ABSTRACT

The bovine genome is known to contain two main repetitive sequence types. Tandemly repeated satellite sequences represent about 20% and short interspersed repeated sequences 6-10% of the total DNA. No other major classes of repetitive DNA have been detected in the bovine genome. Hybridization kinetic experiments have indicated about 12 kinetic families of the short interspersed repeated DNA sequences (Mayfield, et al., 1980). Evidence presented here shows that these kinetic families are sequence related.

Bovine DNA was shotgun cloned into pBR322 and 26 clones containing repetitive DNA were examined. Two of these clones, pPGA7 and pPGB9, were studied in detail. Restriction endonuclease maps indicate that the repetitive unit in each clone is limited to a few hundred base pairs. pPGB9 and 21 out of the remaining 25 clones hybridize to pPGA7. Trace amounts of labelled genomic calf thymus DNA were hybridized with excess clone DNA. The results indicate that the sequence family represented in pPGA7 and pPGB9 comprises about 6% of the genome. This suggests that most, if not all, the short interspersed repeated sequences in the bovine genome belong to the same sequence family.

Both clones were subcloned into bacteriophage M13, and partially sequenced using the method developed by Sanger. A subfragment of pPGA7 was also sequenced by the Maxam-Gilbert method. The DNA sequences resulting from both sequencing methods were compared to each other, to the human Alu sequences, and to a bovine repetitious DNA sequence located near

the bovine corticotropin-beta-lipotropin precursor gene (Watanabe, et al., 1982).

INTRODUCTION

The research in this dissertation characterizes a class of repetitive DNA in the bovine genome known as the short interspersed repeated sequences.

The introduction contains a section on the historical background of DNA sequence organization, the definition of the types of repetitious DNA found in higher eukaryotes, the discovery of the Alu family of short interspersed repeated sequences and its possible functions, and a description of the bovine genome. This is followed by a description of the cloning vectors pBR322 and M13, and then concludes with a discussion of the Maxam-Gilbert and 'dideoxy' DNA sequencing techniques.

Historical Background

The genomes of all higher eukaryotes which have been studied contain repeated and nonrepeated DNA sequences. This was not known, however, until the early 1960s. Using DNA hybridization techniques, Hoyer, et al. (1964) showed that hybridization of DNAs from vertebrates not only occurred, but that the reassociation occurred faster than the reassociation of bacterial DNAs. This was an astonishing result for it had been assumed that animal DNAs would reassociate much slower due to their large genome size. The only possible explanation of these results was that instead of all the DNA sequences being unique, there were multiple copies of certain sequences. A study was then undertaken to see if any of this repetitious DNA could be found in other animal genomes. Michael Waring and Roy J. Britten (1966)

showed that mouse satellite DNA reassociated very rapidly and represented one tenth of the DNA in the mouse genome. They also tried reassociating mouse main band DNA and showed that it, too, contained sequences which were repeated.

One of the most intensely studied DNAs in the 1960s was that of the bovine. This DNA was obtained from calf thymus glands which are rich in DNA and readily available from meat packing plants. When calf thymus DNA was reassociated, it was shown to contain two fractions which reassociated at different rates (Britten and Smith, 1971; Britten and Kohne, 1968). The faster reassociating bovine DNA represented about 40% of the genome and reassociated about 60,000 times faster than the rest of the DNA (Britten and Kohne, 1968). Britten and Kohne (1968) also showed that repeated DNA sequences are present in a diversity of organisms including protozoans, mammals, and plants (Britten and Davidson, 1971; Britten et al., 1976; Britten and Kohne, 1970).

The next questions to be asked were what kinds of repeated sequences are there, and how are they organized in the genomic DNA. R. J. Britten (1971) observed that there was a range of thermal stabilities in the reassociated, repetitive DNA hybrids which suggested sequence divergence among hybrids. This thermal stability was used as the criterion for including or excluding hybrids as belonging to a particular kinetic family of repetitive DNA. He also made the observation that large fragments of DNA, about 15Kb, formed large 'networks' when the DNA strands were allowed to reassociate even for only short periods of time. From this, Britten (Britten and Smith, 1971) concluded that most long pieces of DNA contained

repeated sequences. Since then, many researchers have shown that repetitive DNA sequences are interspersed with varying lengths of nonrepetitive DNA sequences (Davidson, et al.,1973; Davidson, et al., 1975; Manning, et al., 1975; Deininger and Schmid, 1976). Also, most of the organisms studied contain some interspersed repeated sequences which are 200-300 base pairs in length (Davidson, et al., 1975).

Classes of Repetitive DNA

The repetitive DNA sequences have been divided into several classes (Davidson, et al., 1975; Lewin, 1977). These classes are as follows: repeated structural genes such as rRNA, tRNA, 5sRNA, and histones; highly repeated sequences also known as satellite or tandemly repeated sequences; long repetitive sequences; inverted repeated sequences also known as snapback sequences; and short interspersed repeated sequences also known as middle repetitive DNA. A brief description of each of these classes follows.

Highly repeated DNA sequences

The highly repetitive sequences seem to be fairly widespread in eukaryotes (Arrighi, et al., 1970; Davidson, et al., 1975). They are usually short sequences of nucleotides, and can be repeated in tandem over a million times. These sequences were originally termed 'satellite' DNA, because, in many cases, an extra band is seen on cesium chloride density gradients. This is due to the 'satellite' DNAs having a different base composition from the main band DNA. There are, however, highly

repetitive, satellite, DNAs which do not have a different base composition from main band DNA, which means that they do not appear as 'satellite' bands on cesium chloride density gradients. This implies that it is the organization into large blocks of repeated sequences and not base composition which is characteristic of this class of repetitive DNA. The highly repeated sequences are usually not transcribed (Flamm, et al., 1969), usually located at the centromere of metaphase chromosomes, and usually found in heterochromatin (Pardue and Gall, 1970; Mayfield and Ellison, 1976).

Long repetitive sequences

Long repetitive DNA sequences are less well-defined. In primates, hybrids of these sequences are thermally unstable, and these sequences reassociate like single copy sequences (Deininger and Schmid, 1979). Recently, a long repeated sequence has been identified 3' to the human beta-globin gene (Adams, et al., 1980). This sequence is about 6.4kb in length, has low hybrid thermal stability, and seems to be repeated about 2000-3000 times in the human genome. Drosophila melanogaster has about 10% of its genome comprised of sequences averaging 5000 nucleotides in length which are repeated 10s or 100s of times (Manning, et al., 1975).

Inverted repeated sequences

Another of the repeated sequence types are the inverted repeated sequences. These 'foldback' or 'snapback' sequences are characterized by palindromes or closely spaced, self-hybridizing sequences on the same strand of DNA (Perlman, et al., 1976). This type of sequence arrangement can be found in all of the other classes of DNA, and in DNA from all

organisms investigated to date.

Short interspersed repeated sequences

The short interspersed or middle repetitive DNA sequences are characterized by having sequence lengths of 200-400 nucleotides which are found interspersed with single copy sequences of several thousand base pairs (Davidson, et al., 1975; Deininger and Schmid, 1976). Most higher eukaryotes have middle repetitive DNA sequences of 200-400 nucleotides. In contrast to tandemly repeated sequences, hybrids of these sequences do not exhibit a great degree of thermal stability. This indicates that there is substantial sequence divergence within families of short interspersed repeated sequences (Houck, et al., 1978; Marx, 1980). Recent studies on short interspersed repeated sequences of human DNA have shown that sequences previously thought to belong to different kinetic families are actually variations of a single concensus sequence (Houck, et al., 1979; Rubin, et al., 1980). This family of short interspersed repeated sequences in the human genome is known as the Alu family.

Alu and Alu-like Short Interspersed Repeated Sequences

The human genome is comprised of rapidly renaturing, middle repetitive, and single copy sequences (Lewin, 1980; Rinehart et al., 1977). About 20-30% of the genome is made up of the repeated sequences (Schmid and Deininger, 1975; Houck, et al., 1978). Approximately 5-7% of the human genome is made up of repeated sequences which are about 300 nucleotides long, and which were originally thought to belong to about 10 kinetic sequence families (Houck, et al., 1979; Jelinek, et al., 1980).

Recently, it has been discovered that about half of the 300 base pair repetitious sequences are cleaved into 170bp and 120bp fragments by the restriction endonuclease Alu I (Houck, et al.,1979; Jelinek, et al.. 1980). This suggested that a majority of the repeated sequences actually belonged to one and not ten sequence families. Rubin et al. (1980) subjected a mixture of the 170bp Alu I restriction digest fragments to Maxam-Gilbert DNA sequencing and were able to determine a consensus sequence. This result meant that the 170bp fragments were all closely related. This result also suggested that sequence mismatching slowed down the DNA hybridization reaction more than had been realized and led to substantial error in the kinetic estimate of the number of repeated sequence families.

Individual clones of the Alu family were prepared by adding Bam HI linkers to the 300bp fragments resulting from S1 nuclease digestion of repetitive sequence hybrids, and cloning into the Bam HI site of the plasmid pBR322. Individual members of these clones, known as the BLUR clones (Bam Linked Ubiquitous Repeat), were then sequenced (Jelinek, et al., 1980; Rubin, et al., 1980; Deininger, et al., 1981). When the sequences of the various BLUR clones are compared to each other and to the consensus sequence, several interesting observations can be made. Of the 12 BLUR clones which have been sequenced, all of them are imperfect dimers of two 130bp fragments arranged in a head to tail arrangement with a 31bp insert in the right half of the dimer sequence. Sometimes Alu sequences are flanked by several short, direct repeats of AT-rich sequences. These short flanking repeats are reminiscent of transposable elements which have

been characterized for yeast and Drosophila (Cameron, et al., 1979; Strobel, et al., 1979). There is a 70% sequence homology between the right and left halves of the dimer, and the right half is more conserved between different BLUR clones than the left. The individual BLUR clones vary from the concensus sequence by an average of 12.8% with 15-20% mismatching between members. (Deininger, et al., 1981; Schmid and Jelinek, 1982; Jelinek and Schmid, 1982).

A surprising result of recent studies is that several other organisms have repeated sequence families that are sequence related to the human Alu family. Alu or Alu-like sequences have been found in the mouse, rat, Chinese hamster, chicken, African green monkey, hnRNA from HeLa cells, avian sarcoma virus, several low molecular weight RNAs, and short sequences near the origin of replication of SV 40 and papovavirus (Jelinek and Schmid, 1982; Jelinek and Schmid, 1982; Deininger, et al., 1981; Grimaldi, et al., 1981). The Alu-like sequence in rodents is 129bp for the mouse BI sequence and 134 for the Chinese hamster type I sequence. These show close homology to the left half of the human Alu sequence (Schmid and Jelinek, 1982; Jelinek and Schmid, 1982). There is another Alu-like sequence found in rats and Chinese hamsters known as the type II Alu-equivalent. These sequences have about 61 nucleotides in common with the rodent type I and the human Alu sequences. The homology then breaks down and the following 96 nucleotides are not obviously related to any other part of the Alu sequence (Jelinek and Schmid, 1982; Balmain, et al., 1982; Reddy, et al., 1981a and 1981b; Harada and Kato, 1980).

Function of Alu Sequences

Long before the discovery of the human Alu family of short interspersed repeated sequences, there was speculation as to the function of interspersed repetitious DNA (Britten and Davidson, 1969; Davidson and Britten, 1979). Others felt that these repetitious DNA sequences served no particular function and were 'selfish' DNA (Orgel and Crick, 1980; Doolittle and Sapienza, 1980). With the discovery of the Alu family and its representation in a variety of other organisms, the effort to attribute a function to these sequences has intensified.

The presence of portions of the Alu family in double stranded hnRNA of HeLa cells has raised the possibility that these sequences play some role in processing the hnRNA, or a role in transport of mRNA through the nuclear membrane (Jelinek and Schmid, 1982).

As mentioned, there is homology between sequences near origins of replication of some viruses and the Alu sequence. This suggests that the Alu family might be origins of replication in higher organisms. Taylor and Watanabe (1981) have shown that the DNA in CHO cells may contain 250,000 replication origins and they estimate that this is in close accord with the number of Alu family members in the CHO genome.

It has also been shown that the Alu family members of the human genome can be transcribed into small RNAs *in vitro* by RNA polymerase III (Duncan, et al., 1981). The RNA polymerase III *in vitro* transcripts of the Alu sequences are of a specific length for each clone. The 5' ends of all these transcripts are the same but there are different sequences at the 3' ends (Schmid and Jelinek, 1982). The rodent type II Alu sequences

are also transcribed by RNA polymerase III, and similar transcripts in the Chinese hamster have now been found to exist *in vivo* (Haynes and Jelinek, 1981).

Some Alu sequences have a structure similar to transposable elements (Haynes, et al., 1981). These repeated sequences could be remnants of, or still function as transposable elements. Two characteristics of transposable elements are, their ability to form small extrachromosomal circular DNAs, and to induce DNA sequence polymorphisms (Flavell and Ish-Horowicz, 1981; Strobel, et al., 1979). Calabretta, et al., (1982) isolated a human DNA clone which contained a cluster of at least 10 Alu sequences. This clone cross-hybridized to both BLUR 2 and BLUR 8 under conditions for which BLUR 2 and BLUR 8 do not cross-hybridize. A region of this DNA fragment, which contained part of a single copy region flanked by Alu repeats, was subcloned and hybridized to DNA from normal human cells and from a patient with acute lymphocytic leukemia. They found that there were not only restriction endonuclease fragment polymorphisms, but that there were about 5 copies of this sequence in the diseased cells and about 50 copies found in the normal cells. DNA from these same cells was then hybridized with a clone of human placental lactogen, and they found that both diseased and normal cells had the same copy number for this single copy sequence. They concluded that the Alu sequences, like transposable elements, might induce polymorphisms in the human genome. Whether the polymorphism seen in the acute lymphocytic leukemia cells, was a cause or a result of the particular disease is not clear. Calabretta, et al. (1982) also noticed a 4.8kb DNA restriction fragment which

hybridized to the human DNA clone. This fragment could be isolated as circular DNAs of 4.8Kb in length. Krolewski, et al. (1982) have shown the presence of small polydisperse circular (spc) DNAs in an African green monkey kidney cell line. They stated that these are not mitochondrial DNAs, and that about 1% of the spcDNAs in the cells are homologous to the monkey Alu sequence. The entire Alu sequence minus the flanking repeats is seen in these spcDNAs.

About one-half of polysomal poly(A) RNA reassociates at low Rot to form branched structures as seen in the electron microscope (Calabretta, et al., 1981). The mRNAs were shown to hybridize to human Alu clones. The Alu sequences hybridized to one terminus (13%), the middle (53%), or to low molecular weight mRNAs (35%). One such low molecular weight RNA from human cells, now called 7S RNA, has been characterized by two different groups (Ullu and Melli, 1982; Walter and Blobel, 1982). The 7S RNA is found in the cytoplasm of cells and is homologous to the 5′ but not the 3′ end of the human Alu sequence (Ullu and Melli, 1982). This RNA has recently been shown to be associated with the signal recognition protein (Walter and Blobel, 1982). The signal recognition protein is the adapter between the cytoplasmic translation machinery and the translocation machinery in the membrane of the endoplasmic reticulum (Lewin, 1982). Walter and Blobel (1982) have shown that this RNA is essential for the translocation of protein across the endoplasmic reticular membrane during protein synthesis. This is the first definite evidence of a specific function for a DNA or RNA which contains a portion of the Alu family sequence.

The Bovine Genome

In some respects, the bovine genome is similar to the human genome. The bovine short interspersed repeated sequences comprise about 5-10% of the genome, are 300-400 nucleotides in length, and are divided into 8-14 kinetic families which are repeated 50,000-60,000 times (Mayfield, et al., 1980). Watanabe, et al. (1982) have found a family of repetitive sequences in bovine DNA which are only 120 nucleotides in length and estimate that this sequence is present in about 100,000 copies. This sequence does not exhibit extensive sequence homology with the human Alu or other Alu-type sequences; however, it does contain a short sequence which is homologous to sequences at or near the origin of replication for human BK virus, SV 40, and polyoma virus.

pBR322

pBR322 is an engineered plasmid which uses *E. coli* strain K12 as its host bacterium. The plasmid exists in the bacteria as a double stranded circular DNA molecule with length of 4362 base pairs. The plasmid contains the genes for tetracycline and ampicillin resistance which are used for selection of recombinant clones (Cohen, et al., 1969). pBR322 undergoes a relaxed mode of replication, cannot be transferred by conjugation, and can be amplified from around 20 copies/bacterium to 200-1000 copies/bacterium following addition of chloramphenicol (Bolivar, et al., 1977a, 1977b; Lewin, 1977).

The genes for tetracycline and ampicillin resistance originated from

a cotransformation of E. coli with plasmid pSC101 (Tc^r) and the transposon TnA of plasmid pRSF2124 (Ap^r). One resulting recombinant plasmid from this cotransformation was called pBR312. By selective deletion, a plasmid called pBR313 was developed which contained the Tc^r gene and the Ap^r gene which could no longer be translocated (Bolivar, et al., 1977a). Selective deletion of pBR313 resulted in two new plasmids called pBR318 and pBR320. These two plasmids were digested with restriction endonucleases to give a combined total of 5 fragments. The fragments were ligated and one resultant plasmid contained DNA from both pBR318 and pBR320 the Tc^r and Ap^r genes, the origin of replication and useful restriction endonuclease sites. This new plasmid was pBR322 (Bolivar, et al., 1977b).

M13

M13 is a filamentous bacteriophage with a capsid consisting of one major protein type. This capsid surrounds a single stranded circular DNA molecule which is folded back on itself (Lewin, 1977). M13 is a male specific phage which requires an F^+ bacterium. M13 infects, but does not kill the host bacterium.

Upon infection, the positive sense, single stranded DNA of the virus is immediately converted to a double stranded replictive form (RF) by using an RNA primer and host DNA polymerase III (Lewin, 1977). Once the RF form is established, there are two pathways which the viral DNA follows. One pathway produces more RF molecules and the other produces the mature virus. The ability of the RF molecule to enter one of the two pathways seems to be dependent on the viral products of genes II and V.

If there is not much V protein present, then protein II seems to become involved in the production of more RF molecules. When protein V is present in high concentrations, it complexes to the newly forming positive strands preventing the synthesis of the complimentary negative strand. The viral DNA complexed to protein V is circularized and the protein V is replaced by the major capsid protein, VIII. This assembly takes place at the inner cell membrane of the bacterium (Lewin, 1977).

A bacterium which has the M13 virus but has lost the F episome can still produce mature virus particles. However, if the bacterium has lost the F episome, it can no longer be infected by M13. When M13 infects a bacterium, the bacterial growth rate is slowed. This results in 'turbid' plaque formation on a lawn of E. coli (BRL, 1980). These are not 'true' plaques which are devoid of living cells, but are areas of lower bacterial density due to the slower growth rate.

Unlike viruses such as lambda, all the genes in the M13 genome are essential, so deletion of or insertion into any gene inactivates the virus. However, Gronenborn and Messing (1977) found a small region between genes V and II which would accept insertion of foreign DNA (see Appendix A). They subsequently inserted a Hind II fragment from the E. coli Lac operon which would undergo alpha-complimentation with a portion of the Lac operon which was located on the F episome (Gronenborn and Messing, 1977; BRL, 1980). The portion of the Lac operon inserted into the virus contained the I' gene, operator, promotor, and first 145 amino acids of the beta-galactosidase gene (Z gene) (Gronenborn and Messing, 1977; BRL, 1980). This allowed easy selection of recombinant clones,

because a bacterium with the Lac operon containing F episome, and the new M13 virus called M13mp1 would give blue plaques under the appropriate conditions, whereas a bacterium containing M13mp1 which had foreign DNA inserted into the Lac region, would give white plaques. The mechanism for the production of colored plaques will be discussed in the next section.

In order for M13mp1 to be a more useful cloning vector, restriction enzyme sites needed to be located in the DNA of the coding region for beta-galactosidase. This was accomplished by treating M13mp1 with N-[^{3}H] methyl-N-nitrosourea which methylates guanines at the O^{6} position (Gronenborn and Messing, 1978). The single stranded viral DNA was treated with low levels of this compound and a new M13 vector called M13mp2 was isolated. M13mp2 had a new Eco RI site in the DNA sequence coding for the fourth, fifth, and sixth amino acids of beta-galactosidase. The sequence ACG GAT TCA found in mp1 had been changed to ACG **AAT** TCA. The boldface type indicates the new Eco RI site (Gronenborn and Messing, 1978). The new vector, M13mp2, still produced a functional beta-galactosidase polypeptide fragment, and blue plaques formed on indicator plates. However, upon insertion of foreign DNA into the Eco RI site of M13mp2, the beta-galactosidase polypeptide fragment was no longer functional, and white plaques were observed

In order to make the M13mp2 system even more flexible for cloning, DNA fragments containing several restriction endonuclease sites were synthesized *in vitro*, and inserted into the Eco RI site of M13mp2. The restriction sites were chosen so that they would occur in M13 only once. The newly constructed phages still produced a functional beta-

galactosidase fragment unless foreign DNA was inserted into one of the new restriction sites. The first such vector was M13mp7 (BRL, 1980), followed by M13mp8, and mp9 (see Appendix A). Most of the M13 subcloning in this study used M13 mp8.

Plaque formation

As noted in the previous section, the M13 bacteriophage produces 'turbid' plaques which appear blue or white depending on whether a piece of foreign DNA has been inserted into the cloning region of the virus. This color selection is based on an insertional inactivation of the beta-galactosidase gene. This color selection depends on mutations in a strain of E. coli K12 (JM103) as well as mutations in the F episome. This section will first describe these modifications separately and then describe how they interact to produce the color selection.

JM103 is a modified E. coli K12 strain which is proline deficient. Also, the Lac promotor region on the bacterial chromosome has been deleted (BRL, 1980).

The F episome compliments the bacterial proline deficiency. It also contains the Lac operon with a deletion of part of the structural gene for beta-galactosidase. In addition to this deletion, there is a deletion in the Lac repressor gene (I gene) which leads to overproduction of lac repressor (BRL, 1980).

The M13 virus contains the I gene, operator, promotor, and part of the beta-galactosidase gene in the cloning region. The engineered restriction endonuclease sites used for cloning are located within the beta-galactosidase gene element (BRL, 1980).

The complimentation of the bacterial proline deficiency by the F episome provides a selective mechanism for the retention of the F episome. A JM103 bacterium which contains the F episome is still lac^- because of the partial deletion of the episomal Z gene. If an M13 virus of the mp series infects this cell the bacterium is then lac^+, because the F episome and the virus produce different functional fragments of beta-galactosidase. These different functional fragments are able to combine in the cytoplasm of the bacterium to produce a functional beta-galactosidase molecule. The overproduction of lac repressor means that all the Lac operators have bound repressor, so eventhough the genotype is lac^+, the phenotype is lac^-.

To activate the Lac operator, the nonmetabolized inducer isopropyl-beta-D-thiogalactopyranoside or IPTG is added to the medium. IPTG binds to the lac repressor molecule releasing it from the operator so transcription of beta-galactosidase will take place. An exogenous substrate for beta-galactosidase is also added to the medium. This is a lactose analog 5-bromo, 4-chloro, 3-indolyl-beta-D-galactoside or Xgal. When Xgal is metabolized, a blue color is produced.

The JM103 bacterium, containing the F episome and infected with an M13 virus, grown in the presence of IPTG produces a functional beta-galactosidase, and Xgal will be metabolized to produce a blue plaque. However, if the M13 virus contains an insertion of DNA, the beta-galactosidase will not be functional, Xgal cannot be metabolized, and a white plaque results (BRL, 1980).

DNA Sequencing

Two methods of DNA sequencing were used in this project: 'dideoxy' and Maxam-Gilbert.

Dideoxy DNA sequencing

Dideoxy DNA sequencing was developed by Sanger, et al. (1977). This makes use of DNA polymerase I's inability to extend a DNA chain which terminates in a 2',3'-dideoxynucleotide. This procedure utilizes the Klenow fragment of DNA polymerase I which lacks the 3'-5' exonuclease activity (Jacobsen, et al., 1974), but which can still synthesize a complimentary copy of a single stranded DNA template provided it is primed with a short double stranded DNA region.

The M13 cloning vectors are ideal for this sequencing method (Messing, et al., 1981). The virus is single stranded DNA which means that a large quantity of pure template can be easily obtained. All the cloning sites are clustered in one small region of the viral DNA, so that a single universal primer can be used to sequence 200-300 nucleotides into each cloned piece of DNA. Prior to the development of the M13 cloning vectors, a specific primer had to be made for every 200 base pairs sequenced.

Four reactions are routinely set up, each containing the four deoxynucleoside triphosphates (one of which is labelled with ^{32}P) and one of the dideoxynucleoside triphosophates (BRL, 1980; Grossman and Moldave, 1980, p.499). The concentration of the deoxy and dideoxynucleoside triphosphates is such that a dideoxynucleoside triphosphate will be

randomly incorporated once per 100-200 nucleotides at every position in the complimentary strand where that base appears. The four reaction mixtures are electrophoresed on a thin polyacrylamide gel and the sequence visualized by autoradiography. Since the primer is located at the 5′ end of the sequenced strand of DNA, the sequence read from the bottom to the top of the gel is in the 5′-3′ direction.

Maxam-Gilbert DNA sequencing

The Maxam-Gilbert DNA sequencing method (Maxam and Gilbert, 1977) is based on chemical removal of the DNA bases followed by chain breakage. In this method, the DNA to be sequenced can be any length, although only 200-300 nucleotides from the labelled end can be determined at one time. This method can use single or double stranded DNA as the starting material with the stipulation that only one strand and only one end of that strand is labelled with ^{32}P.

The bases in the DNA are modified using hydrazine to modify C′s and T′s, hydrazine plus 1.5M NaCl to modify only C′s, formic acid plus piperidine to modify A′s and G′s, and dimethyl sulfate to modify only G′s. After base modification and removal of the modified base, the phosphate backbone of the DNA chain is cleaved at the modification by piperidine (Grossman and Moldave, 1980, p.499). The mixtures of DNA fragments cleaved in various positions are electrophoresed on the same type gels as the dideoxy sequencing method. However, the autoradiographic band pattern is different. To read the Maxam-Gilbert sequences, the lane containing the reactions for modifying G′s and A′s is compared with that in which only G′s are modified. The C and T lane and the C lane are also compared.

Bands which appear in both the compared lanes are G's or C's, and those which are only in one of the lanes are T's or A's. The other factor which has to be taken into consideration when reading a Maxam-Gilbert gel is whether the 3' or 5' end of the sequenced DNA is labelled. This determines whether the sequence read from the bottom to the top of the gel is in the 3'-5' or 5'-3' direction.

MATERIALS AND METHODS

This section describes the procedures, and the materials used in all experiments. Solutions and buffers used in this section are defined in Appendix B. Drawings of some of the apparatus used can be found in Appendix C.

Bacterial Strains, Viral Strains, Plasmids, and Growth Conditions

E. coli HB101 (r^-K, m^-K, dam^+, dcm^+) was used as the host bacterial strain for the plasmid pBR322 (Bolivar, et al., 1977b) and pBR322 derived recombinant plasmids. HB101 was grown in LB medium at 37°C with vigorous shaking. HB101 cells containing plasmid were grown in LB medium plus 200 ugs/ml ampicillin.

E. coli JM103 (Δlacpro, thi, strA, endA, sbcB15, hsdR4, supE, F'traD36, proAB, $lacI^q$, ZΔM15) was used as the host strain for bacteriophages M13mp7, M13mp8, and M13mp9 (BRL, 1980). The JM103 cells were grown in YT medium at 37°C with vigorous shaking. The JM103 cells were stored in 50% glycerol at -70°C and periodically replated on YT minimal agar plates to insure the retention of the F episome.

Phage stocks were assayed by a plaque assay method (BRL, 1980). YT medium was inoculated with a fresh loop of JM103 and this was allowed to reach a cell density of $5x10^8$ cells/ml. 0.3 mls of these exponentially growing cells were either mixed with 0.1 ml of 100-fold serial dilutions of the phage stock, or the total transformation mixture. 10 uls of 100 mM

IPTG in water and 35 uls of 2% Xgal in deionized formamide were added to each tube of either transformed JM103 or JM103 plus phage. 4.0 mls of YT soft agar at 42°C were added to each tube, the contents mixed gently, and the entire mixture poured onto YT agar plates and incubated overnight at 37°C.

When assaying for transformation, control plates containing just the JM103 cells (check for contaminating virus), and those containing 2 ngs of M13mp8 RF (check for transformation efficiency) were run in parallel. Blue plaques on plates indicated undigested or religated M13mp8, while white plaques indicated a deletion or insertion of DNA into the beta-galactosidase gene.

When assaying for phage concentration, a control plate with just JM103 was poured along with the serial dilution plates. Phage concentrations/ml were calculated by taking the number of blue plaques on a plate, multiplying that number by 10, and correcting for the appropriate dilution.

A small collection of recombinant plasmids was used as a source of specific repetitive bovine DNA sequences. These recombinant plasmids had been previously created by Peter Good in our laboratory by the shotgun cloning of Bam HI fragments of bovine DNA into the Bam HI site of pBR322. Selection for the presence of repetitive DNA sequences in these clones was by colony hybridization with tritium labelled total calf thymus DNA. These plasmids have been designated as the pPG series (pPGA7, etc.) and the plasmids pPGA7 and pPGB9 were selected for further study.

Purification of Calf Thymus DNA

Bovine DNA was prepared from frozen calf thymus which had previously been obtained from the Iowa State University meat lab. 10.0 grams of calf thymus and 120 mls of saline-EDTA were homogenized in the cold using a Waring Blender. The blender was connected to a variable transformer and run at 85V for 30 seconds and then 20V for 3 minutes. The homogenate was strained through two layers of Miracloth (Calbiochem), put into 50 ml plastic centrifuge tubes, and centrifuged in the cold for 10 minutes at 3500rpm in an IEC model HRI centrifuge.

Each pellet was resuspended, using a metal spatula, in 15-25 mls of cold saline-EDTA, and if necessary, homogenized gently with a glass-teflon homogenizer to thoroughly resuspend the pellets. The homogenate was then centrifuged in the cold for another 10 minutes at 3500rpm. Each pellet was resuspended in 20 mls of cold 10 mM TRIS (pH8) and adjusted to 0.5 M NaCl and 2% SDS. The resulting viscous solutions were then diluted with 10 mM TRIS (pH8), 0.5 M NaCl, 2% SDS until they were viscous but pourable (about 120 mls final volume).

An equal volume of redistilled, water saturated phenol was added, the mixture shaken in the cold for 15 minutes, transferred to 250 mls glass centrifuge bottles and centrifuged in the cold for 5 minutes at 2000rpm. The aqueous DNA containing layer (top layer) was removed with a large pipet. If the aqueous layer could not be removed without the phenol layer mixing in, then it was diluted with additional 10 mM TRIS (pH8), 0.5 M NaCl, 2% SDS and recentrifuged. This aqueous layer was treated successively with 10 ugs/ml of RNAse A (SIGMA) for 45 minutes at 37^{o}C and

for 2 hours at 37°C with 100 ugs/ml of pronase, which had been previously autodigested for 30 minutes at 37°C. The DNA solution was phenol extracted as before and treated again with 100 ugs/ml pronase for 2 hours at 37°C, followed by two additional phenol extractions.

After the last phenol extraction, 2 volumes of ice cold 95% ethanol were added to the aqueous layer and the DNA was spooled onto a glass rod. The rod plus DNA was put into a glass screw cap tube containing 5-10 mls of 10 mM TRIS (pH8), 1 mM EDTA (TE) solution, and the DNA was allowed to dissolve overnight at 4°C. The dissolved calf thymus DNA was dialyzed against two, one liter changes of 10 mM TRIS (pH8), 1 mM EDTA.

The purity of the DNA was determined by measuring absorbances at 230_{nm}, 260_{nm}, and 280_{nm} using a Beckman model DU spectrophotometer. The DNA was considered to be adequately purified if the A_{260}/A_{280} was greater than 1.8 and the A_{260}/A_{230} was between 0.4 and 0.5. If the ratios were not close to these values, the DNA was pronased, phenol extracted, ethanol precipitated, redissolved, and the absorbances checked again.

Radioactively Labelled Bovine DNA

Tritiated bovine DNA was purified from bovine cultured cells grown in the presence of ^3H thymidine. 180 mls of cultured bovine kidney cells at $1.2x10^8$cells/ml in MEM-Eagles medium (minus thymidine), supplemented with 10% fetal calf serum, penicillin and streptomycin, were obtained from Dr. Gerald Woode, Department of Veterinary Microbiology, Iowa State University. The cells were divided among nine $75cm^2$ plastic culture flasks and incubated with 55uCi of ^3H-thymidine (New England Nuclear, SA

2 Ci/mMol) at 37°C with CO_2 and humidity. The flasks were allowed to become confluent, which took 48 hours.

The medium was poured off and each bottle washed with 3.0 mls of a trypsin solution. 5.0 mls of the trypsin solution were then added to each bottle and left until the cells started to detatch. The action of trypsin was stopped by adding 3.0 mls of calf serum and 3.0 mls of MEM medium to each bottle. A sterile rubber policeman was used to scrape cells off the bottles and the cells were centrifuged at 1000rpm for 10 minutes. The supernatants were poured off and each cell pellet resuspended in 1.0 ml of cold TNKM. All cells were then combined into a single centrifuge tube and pelleted again at 1000rpm for 10 minutes in the cold.

The resulting pellet was resuspended in 4.5 mls cold TNKM and 0.5 mls of 10% TRITON X100, homogenized gently with a glass-teflon homogenizer, centrifuged at 5000rpm for 10 minutes at 4°C, and the pellet resuspended in 8.0 mls of cold 0.01 M TRIS-HCl (pH8). 1.0 ml of 5.0 M NaCl and 1.0 ml of 20% SDS were then added and mixed well until the mixture became very viscous. The homogeneous solution was then incubated with 10 ugs/ml RNAse for 45 minutes at 37°C followed by 100 ugs/ml of preincubated pronase for 2 hours at 37°C. This solution was phenol extracted two times and the aqueous layer dialyzed against two, one liter changes of 10 mM TRIS(pH8), 1.0 mM EDTA. The dialyzed DNA solution was pronased, phenol extracted, and dialyzed again. A few drops of chloroform were then added and the DNA solution was stored at 4°C. The specific activity of the DNA was 2.6×10^6 cpm/ugDNA.

Purification of Plasmid and Viral RF DNA

Quicky lysate procedure

For small quick preps of plasmid or viral RF DNAs, a 'quicky' lysis procedure was adapted from Birnboim and Doly (1979).

Plasmid-containing or M13-infected bacterial cells were grown overnight in 2.0 mls of LB or YT medium at 37°C with shaking. The cultures were poured into sterile, 1.5 ml Eppendorf tubes and centrifuged in the cold for 2 minutes in the microfuge. The supernatants were discarded and the cell pellets resuspended in 100 uls of 2 mgs/ml egg white lysozyme (SIGMA) in sterile 50 mM glucose, 10 mM EDTA, 25 mM TRIS (pH8). The solutions were placed on ice for 30 minutes, then 200 uls of 0.2N NaOH - 1% SDS were added to each tube. The tubes were vortexed and left on ice for an additional 5 minutes.

After 5 minutes, 150 uls of 3.0 M sodium acetate (pH4.8) were added, the tubes were vortexed, and then placed on ice for 1 hour. Following this incubation, the tubes were centrifuged in the microfuge at 5°C for 10-15 minutes. 0.4 mls of each supernatant were transferred to another sterile 1.5 ml Eppendorf tube, 1.0 ml of ice cold 100% ethanol was then added, and the DNA precipitated at -70°C for at least 30 minutes. The precipitates were collected by a 10-15 minute centrifugation in the microfuge.

The precipitates were dissolved in 100 uls of 0.1 M sodium acetate (pH8). An equal volume of redistilled phenol was added, the tubes vortexed, put on ice for 5 minutes, vortexed again, and then centrifuged for 1 minute in the microfuge. The upper aqueous phase was transferred

to another sterile Eppendorf tube, two volumes of 100% ice cold ethanol added, and the DNA precipitated as previously described. The pellets were washed with 1.0 ml of ice cold 70% ethanol, centrifuged for 5 minutes, and the pellets dried on the lyopholyzer for 5 minutes.

For restriction enzyme digests, the dried DNA pellets were resuspended in 30 uls 10 mM TRIS (pH8), 1 mM EDTA, and 10 uls of this were used for each digestion and gel well.

Cleared lysate procedure

For large preps of plasmid or viral RF DNA, a 'cleared' lysis procedure was performed (Clewell and Helinski, 1970; Clewell, 1972) followed by a cesium chloride (CsCl) density gradient (Meselson, et al., 1957).

500 ml cultures were grown overnight at room temperature in 2 liter flasks with vigorous aeration. If recombinant plasmids were being grown, they were amplified by addition of 200 ugs/ml chloramphenicol when the cell density reached $5x10^8$ cells/ml. After growth, the cultures were poured into 250 ml plastic screw cap centrifuge bottles and centrifuged at 5000rpm for 10 minutes in a Sorvall GSA rotor. Each cell pellet was resuspended in 40-50 mls of cold 0.25 M TRIS, 0.25 M EDTA, 0.1 M NaCl (pH8) and recentrifuged as before. These pellets were resuspended in 10-15 mls of 25% sucrose, 50 mM TRIS-HCl (pH8), 3.0 mgs/ml lysozyme and incubated at room temperature for 1 hour. Following incubation, 2.0 mls of 0.25 M TRIS-HCl, 0.25 M EDTA (pH8) were added, and 0.2 mls of TRITON X100. If lysis did not occur, i.e. the solution did not become translucent and very viscous, 5.0 M NaCl was added to a final

concentration of 0.5 M.

The solutions were then poured into plastic screw cap tubes and centrifuged in a Beckman 30 rotor at 25,000rpm for 2 hours at 15°C. The supernatants were carefully poured into 30 ml Corex tubes (at this stage the pellets were loosely attached to the centrifuge tube wall, so care was taken when pouring off the supernatant). The volume of each supernatant was measured. 1.0 gram of CsCl (KBI technical grade) and 0.1 ml of 10.0 mgs/ml ethidium bromide were added to each 1.0 ml of DNA solution. A filler solution of 1.0g CsCl, 0.1 ml ethidium bromide, and 1.0 ml TE was used to make the final volume of each DNA-CsCl solution 16.5 mls.

The DNA-CsCl solutions were transferred to 17 ml polyallomer centrifuge tubes (Sorvall) and centrifuged in the Sorvall TV865B verticle rotor for at least 18 hours at 15°C and 45,000rpm. The reograd mode of the Sorvall OTD-50 ultracentrifuge was used when decelerating the run. A hemostat was used to carefully remove the tubes from the rotor. The tubes were then placed in a Hoeffer gradient collection apparatus. Long wave UV light from a hand held mineral light (U.V. Products Inc.) was used to visualize the DNA bands in the tubes. Generally, two bands were visible. A 3.0 ml disposable syringe with a 20-gauge needle was used to remove the lower band. The needle was inserted through the side of the tube about 1/8 inch below the band, and CsCl solution was withdrawn until the DNA band was completely collected.

An equal volume of 1-butanol was then mixed with the DNA-CsCl solution in a glass screw cap test tube. The phases were allowed to separate, the top phase (butanol, which was pink due to the ethidium

bromide) was removed and discarded. Butanol extractions were repeated until the CsCl began to precipitate. This step not only removed the ethidium bromide, but concentrated the DNA. The DNA solution plus precipitated CsCl was then dialyzed overnight against one liter of TE plus 0.5 M NaCl. The dialyzed solution was then phenol extracted, and dialyzed twice against one liter of TE.

Growth and Purification of M13 Phage and M13 Phage DNA

5.0 mls of sterile YT medium were inoculated with a loop of fresh JM103 and 50 uls of M13 or recombinant M13 virus stock, and incubated for 7 hours at 37°C with vigorous shaking. The culture was transferred to a sterile 15 ml Corex centrifuge tube, and the bacteria pelleted by a 20 minute centrifugation at 10,000rpm. The supernatant was carefully poured into another sterile 15 ml Corex tube and recentrifuged for an additional 20 minutes. The supernatant was transferred to a sterile, plastic disposable screw cap test tube and 1.5 mls of sterile 27% polyethylene glycol (PEG) 6000, 3.3 M NaCl were added to each tube. The tubes were mixed well and incubated in an ice bucket in the refrigerator for 3-4 hours.

The PEG-precipitated M13 virus was pelleted by 5 successive 5 minute centrifugations in sterile 1.5 ml Eppendorf tubes. After the final centrifugation, the inside of each Eppendorf tube was wiped with a Kimwipe to remove any residual PEG, which interferes with DNA sequencing reactions. The viral pellet was resuspended in 100 uls of sterile TE and stored at 5°C.

If single stranded, viral DNA was needed, the pellet was resuspended in 100 uls of sterile TE, and 100 uls of redistilled phenol were added. This mixture was vortexed, put on ice for 5 minutes, vortexed again, and centrifuged in the microfuge in the cold for 5 minutes. The upper aqueous layer was transferred to another sterile Eppendorf tube using a 100 ul capillary pipette. 50 uls of redistilled phenol and 50 uls of chloroform were then added, the mixture vortexed, iced, vortexed, and centrifuged as before. The upper aqueous layer was again transferred to another Eppendorf tube and subsequently phenol-chloroform and chloroform extracted. One-tenth volume of 3.0 M sodium acetate (pH8) was added to the aqueous phase, then 250 uls of ice cold 100% ethanol were added to each tube and the DNA precipitated at $-70^{o}C$ for 30 minutes. The DNA was pelleted by a 10-15 minute centrifugation in the microfuge. Each DNA pellet was washed with 1.0 ml of ice cold 70% ethanol and dried for 5 minutes on the lyopholyzer. The dried viral DNA was stored at $-20^{o}C$.

Transformation of HB101

E. coli strain HB101 was grown in 50 mls of sterile LB medium in a 250 ml Nephelo flask (Bellco Biologicals) at $37^{o}C$ with shaking. The cells were grown until the absorbance at 550 nanometers (A_{550}) was 0.6, which corresponds to about $5x10^{8}$ cells/ml. When an A_{550} of 0.6 was reached, the cells were centrifuged in sterile 40 ml plastic, screw cap centrifuge tubes at 5000rpm for 10 minutes at $5^{o}C$. The supernatant was discarded into a beaker of disinfectant which was later autoclaved.

The cell pellet was resuspended in 1/2 volume of cold 10 mM $MgCl_2$,

1.0 mM TRIS(pH8) buffer. The cells were pelleted as before. The cell pellet was resuspended in 1/10 volume (relative to the original volume of the cell culture) of cold, sterile 50 mM $CaCl_2$ and incubated on ice for 20 minutes. 100 uls of this cell solution were added to cold, sterile 1.5 ml Eppendorf tubes containing 1 ug of plasmid DNA. This mixture was left on ice for 1 hour. After 1 hour, the cells were heat shocked for 2 minutes in a 42°C water bath (Mandel and Higa, 1970). The heat shocked cells were then centrifuged in the Eppendorf microfuge for 2 minutes. Each cell pellet was resuspended in 1.0 ml of sterile LB medium and transferred to sterile 16mm x 150mm culture tubes. These cultures were incubated at 37°C for 2 hours with shaking (Wu, 1979). 10^{-1}, 10^{-2}, and 10^{-3} dilutions of each cell culture were then made using sterile LB medium. 0.1 ml of each dilution was spread onto LB agar plates containing 100 ugs/ml ampicillin. These plates were incubated at 37°C overnight.

For experiments involving recombinant DNA plasmids, a dilution which gave isolated colonies was replica plated onto LB agar plates containing 50 ugs/ml tetracycline. Colonies containing DNA inserted in the tetracycline resistance gene of pBR322 grow on ampicillin plates but not on tetracycline plates. These recombinant plasmids are phenotypically amp^+, tet^-. Colonies containing just pBR322, grow on both plates and are phenotypically amp^+, tet^+.

Individual amp^+, tet^- colonies were picked and grown overnight at 37°C in 2.0 mls of sterile LB medium. Each culture was then stored in 50% glycerol at -70°C. These stored cultures were used as the stock supply for each recombinant plasmid clone. Whenever a particular clone was to be

analyzed, a sterile loop was scraped across the top of the frozen stocks. This was then used to inoculate test tubes or flasks of sterile LB medium, and grown overnight at 37°C with shaking.

Transformation of JM103

Several loops of a JM103 colony grown on minimal agar were added to a sterile 250 ml Erlenmeyer flask containing 30-50 mls of YT medium. This was allowed to grow for 2-3 hours at 37°C with shaking. The cell culture was then poured into sterile, plastic 40 ml screw cap centrifuge tubes and centrifuged at 5000rpm at 4°C for 10 minutes. At this point, the flask that had initially been used for growing the JM103 culture was saved. 10-20 mls of sterile YT medium were added to this flask and the residual bacterial cells in the flask were grown at 37°C with shaking. This culture provided the exponentially growing JM103 needed later in the transformation protocol. The cell pellet, from above, was resuspended in 10-15 mls of cold, sterile 50 mM $CaCl_2$ and left on ice for 20 minutes. This was centrifuged as before and the pellet resuspended in 3-5 mls of cold, sterile 50 mM $CaCl_2$.

0.25 mls of these now competent cells were added to ice cold, sterile 16mm x 150mm culture tubes containing 25 uls of M13 RF DNA (at least 300 ngs of DNA). When the transforming DNA was from a ligation reaction mixture, a volume greater than or equal to one-tenth of the ligated DNA to competent JM103 cells was necessary in order to have efficient transformation. This cell-DNA mixture was left on ice for 40-60 minutes, and heat shocked for 2 minutes in a 42°C water bath. After heat shocking,

the cultures were assayed for transformation by the plaque assay method previously described.

Cloning using M13

The double stranded replicative form (RF) of M13mp8 or M13mp9 was purified on CsCl-ethidium bromide density gradients. For each ligation reaction, 100 ngs of restriction endonuclease digested vector DNA were combined with 500 ngs of enzyme digested DNA to be cloned.

Restriction endonuclease digests were carried out in 30 ul volumes in sterile 1.5 ml Eppendorf centrifuge tubes. 8-12 units of restriction endonuclease were used for each digest. An exception was the enzyme Acc I, where only 1.4 units were used. The digestion mixtures were incubated at 37°C for 1 hour. The restriction enzymes used were: Bam HI, Hinc II, Acc I, Pst I, Alu I, Hae III, Hpa II, Hinc II - Eco RI double digests, and Bam HI - Hind III double digests.

Vector DNAs were routinely treated with bacterial alkaline phosphatase to prevent recircularization of the vector during the ligation reaction. Restriction endonuclease digested DNA was diluted with sterile water to reduce the salt concentration to 50 mM. 30 units of bacterial alkaline phosphatase (BRL) were then added and the mixture incubated at 65°C for 40 minutes. The DNA was spermine precipitated or phenol extracted and ethanol precipitated. DNAs, which were not treated with bacterial alkaline phosphatase, were phenol extracted and ethanol precipitated directly from the restriction digest solutions.

The following recipe was used for 'sticky end' ligation reactions:

5 uls enzyme digested M13mp8 or M13mp9 (100ng)
10 uls enzyme digested clone pPGA7 or pPGB9 (500ng)
5 uls 10 mM ATP
5 uls 10X T4 DNA ligase assay buffer
10 uls T4 DNA ligase (40 units/ul New England Biolabs or 0.02 units/ul BRL)
15 uls sterile distilled water

These ligations were incubated at 14-15^{o}C overnight.

For 'blunt end' ligations the same recipe was used except the ligation reaction mixture was incubated at room temperature overnight.

For mixed 'blunt end' and 'sticky end' ligations the samples were incubated for 8 hours at 14-15^{o}C then overnight at room temperature.

Control ligations containing only enzyme digested M13 vectors were routinely set up in parallel with the cloning ligations.

Following ligation, E. coli JM103 was transformed with one-half of each ligation reaction mixture. The rest of the ligation mixture was stored at -20^{o}C.

Screening M13 Clones

The following procedures are modifications from those found in the BRL manual (1980).

White plaques were picked with sterile toothpicks, transferred to 2.0 mls of sterile YT medium, and grown overnight at 37^{o}C with vigorous shaking. The cells were then poured into sterile 1.5 ml Eppendorf tubes and centrifuged in the microfuge for 3 minutes. Each supernatant was poured into separate sterile 1.5 ml Eppendorf tubes and used as the virus stock for that particular clone. These stocks were kept at 5^{o}C. The

pellets of each clone were resuspended and RF DNA for that clone was extracted by the 'quicky' lysis procedure. A few blue plaques were also picked and treated in the same manner to serve as controls.

DIGE: Direct gel electrophoresis

20 uls of the virus supernatant were transferred to a sterile 1.5 ml Eppendorf tube. 2.0 uls of 10% SDS were added to the 20 uls of viral supernatant and the contents were vortexed. (The SDS solubilizes the viral coat protein releasing the DNA.) 5 uls of the BPB dye mix were then added to the disrupted virus, and the entire sample loaded on a horizontal 0.7% agarose gel in TBE buffer. The gel was electrophoresed at 70V for 10 hours, and then stained with 1.0 ug/ml ethidium bromide in TBE buffer. The viral DNA bands were visualized with UV light, and a picture of the gel was then taken using Kodak Royal Pan film (Estar thick base, 0.18mm).

DNA from any white plaque with an insert of greater than 100bp showed decreased mobility on the agarose gel as compared to the control M13 viral DNA.

C-test: Test for cloning complimentary DNA strands

The viral DNA is always the '+ sense' strand, so recombinant phages containing complimentary DNA strands should hybridize only via their inserts.

20 uls of viral supernatant of each of two different clones were combined and mixed with 4.0 uls of 10% SDS. The mixture was incubated at 65^{o}C for 3 hours. 10 uls of the BPB dye mixture were added to the hybridized DNA and 25 uls of this were loaded onto a horizontal 0.7% or 1% agarose-TBE gel. Electrophoresis was carried out as in the 'DIGE' test.

As a control, one or both of the recombinant viral DNAs used in the hybridization were run in separate wells. Those recombinants, which hybridized, migrated more slowly on the gel than the unhybridized controls.

B-test: Southern blotting

Viral DNAs were electrophoresed and photographed as in the 'DIGE' test. The DNA was transferred to nitrocellulose filter paper or Genatran 45 (D&L Filter Corporation) paper following the Southern transfer procedure (Southern, 1975).

At least 10^6cpms of radioactive, nick translated probe DNA were boiled in 1x Denhardts for 5-10 minutes and then put on ice. The nitrocellulose or Genatran 45 was placed in Daisy Seal-A-Meal bags. The radioactive probe and 10-30 mls of 1x Denhardts plus 0.2% SDS were then added to the bags with the filters. This was heat sealed using the Daisy Seal-A-Meal apparatus. When Genatran 45 was used, the 1x Denhardts was made 0.5% BSA and 1% SDS.

The hybridization was allowed to proceed at 60^{o}C overnight. Each filter sheet was then washed at 50^{o}C for 30 minutes in 250 mls of 2x SSC. This washing step was repeated three more times for each filter sheet. (It was found that gentle rocking of the wash trays resulted in lower background on the autoradiograms.) After the washing was completed, the filter sheets were air dried. The dried filters were then autoradiographed overnight at room temperature using Kodak XAR-5 X-ray film. If necessary, a second piece of X-ray film was placed on the filter sheet, and a second exposure taken for a longer time.

Alternate 'B-test'

Instead of transferring viral DNA from agarose gels to filters, the viral supernatants were sometimes directly spotted onto the filter paper. Two 5 ul aliquots of viral supernatant from recombinants were spotted directly onto nitrocellulose filter paper or Genatran 45 paper. The paper was then baked in a vacuum oven for 2 hours at 80°C. (The baking destroys the viral protein coat and fixes the single stranded viral DNA to the paper.) The filter sheets were then treated and hybridized with a radioactive DNA probe as described under 'B-test'.

RE-test

M13 recombinant RF DNA obtained by the 'quicky' lysis procedure was digested with appropriate restriction endonucleases. The digests were electrophoresed on a horizontal 1% or 2% agarose-TBE gel. To determine the size of the DNA insert, digests of the parent M13 recombinant clone and pPGA7 or pPGB9 were electrophoresed along with standards of known length. By this method, DNA fragments were identified by size only.

Restriction Enzyme Digests for Mapping

Restriction endonuclease digests of 1-3 ugs of DNA were performed in sterile 1.5 ml Eppendorf tubes in a total volume of 20-25 uls. Stock solutions of clone, pPGA7, or pPGB9 DNAs were 0.2-0.3 ugs/ul. 5-10 uls of each stock DNA were used per digestion reaction. Some digests were done using DNA bands electroeluted from 7% acrylamide-TBE gels, precipitated, then resuspended in 10 uls TE. Restriction endonucleases (BRL) were stored in 50% glycerol at -20°C, and at concentrations of 6-10 units/ul.

5-10 units of enzyme were typically used per digest. (Since glycerol can inhibit enzyme digestion of DNA, the final glycerol concentration in the digestion reaction mixture was kept less than 5%.)

The enzyme assay buffers were made as 10 times concentrated sterile solutions and stored at 5°C. The 10x assay buffers were usually those suggested by BRL for that particular enzyme. However, later experiments used 10x assay buffers with low, medium, or high salt concentrations (Maniatis, et al., 1982) instead of 10x buffers for each individual enzyme. A list of restriction endonucleases and whether they use low, medium, or high salt buffers is located in Manaitis, et al. (1982).

Typical single and double enzyme digests were set up as follows:

Single: 10.0 uls DNA stock
2.5 uls 10x enzyme assay buffer
1.0 ul enzyme (5-10 units/ul)
11.5 uls sterile distilled water

Double: both enzymes using the same salt concentrations
10.0 uls DNA stock
2.5 uls 10x enzyme assay buffer
1.0 ul enzyme 1 (5-10 units/ul)
1.0 ul enzyme 2 (5-10 units/ul)
10.5 uls sterile distilled water

Double: enzymes using two different salt concentrations
10.0 uls DNA stock
2.5 uls 10x enzyme assay buffer of lower salt
1.0 ul enzyme 1 (5-10 units/ul)
11.5 uls sterile distilled water
After 1 hour incubation at 37° add:
1.5 uls 10x enzyme assay buffer of higher salt
1.0 ul enzyme 2 (5-10 units/ul)

The digestion mixtures were mixed well and incubated at 37°C for 2 hours to insure complete digestion. After 2 hours, 5 uls of BPB loading dye mix were added to each digestion mixture, and the sample vortexed. The entire sample from each digestion was loaded into one gel well and

electrophoresed on an appropriate percentage agarose-TBE horizontal gel.

Nick Translation

Nick translation (Rigby, et al., 1977) of 2-5 ugs of DNA was performed using a single label [^{32}P]dATP or double labels, [^{32}P]dCTP plus [^{32}P]dGTP, or [^{32}P]dCTP plus [^{32}P]dATP (Ammersham: > 800Ci/mmol). T4 DNA polymerase I from Borhienger-Mannheim was used in all nick translation experiments. This DNA polymerase is the nick translation grade which contains contaminating DNAse.

Figure 1 shows the elution profile of [^{3}H] nick translated DNA from Sephadex G-75, using different amounts of DNA Polymerase I. ^{3}H-nick translated DNA was electrophoresed on 1% agarose, the gel dried, and fluorographed. Figure 2 shows that a vast majority of the radioactivity was contained in large fragments of DNA.

The nick translation reactions were done in a total volume of 60 uls in a 1.5 ml Eppendorf centrifuge tube. Usually, a total of 120uCi of radioactivity was used in each reaction. Sometimes it was necessary to concentrate the radioactivity by lyopholyzation. The standard reaction mixture used was as follows:

2-5 ugs DNA
120 uCi of each [^{32}P]dNTP
2.0 uM of each unlabelled dNTP
4.5 units of DNA polymerase I (Boerhinger-Manheim nick translation grade which contains contaminating DNAse)
6.0 uls of nick translation assay buffer
Sterile distilled water to a total volume of 60 uls

The reactions were allowed to proceed for 2-3 hours at 11-14^{o}C, and

FIGURE 1: Incorporation of ^{3}H dCTP and ^{3}H dGTP into DNA by nick translation

1.0 ug of DNA was incubated with 0 units (x), 1.0 units (o), or 4.5 units (*) of T4 DNA polymerase I. The DNA was fractionated on Sephadex G-75 and the radioactivity in each fraction determined.

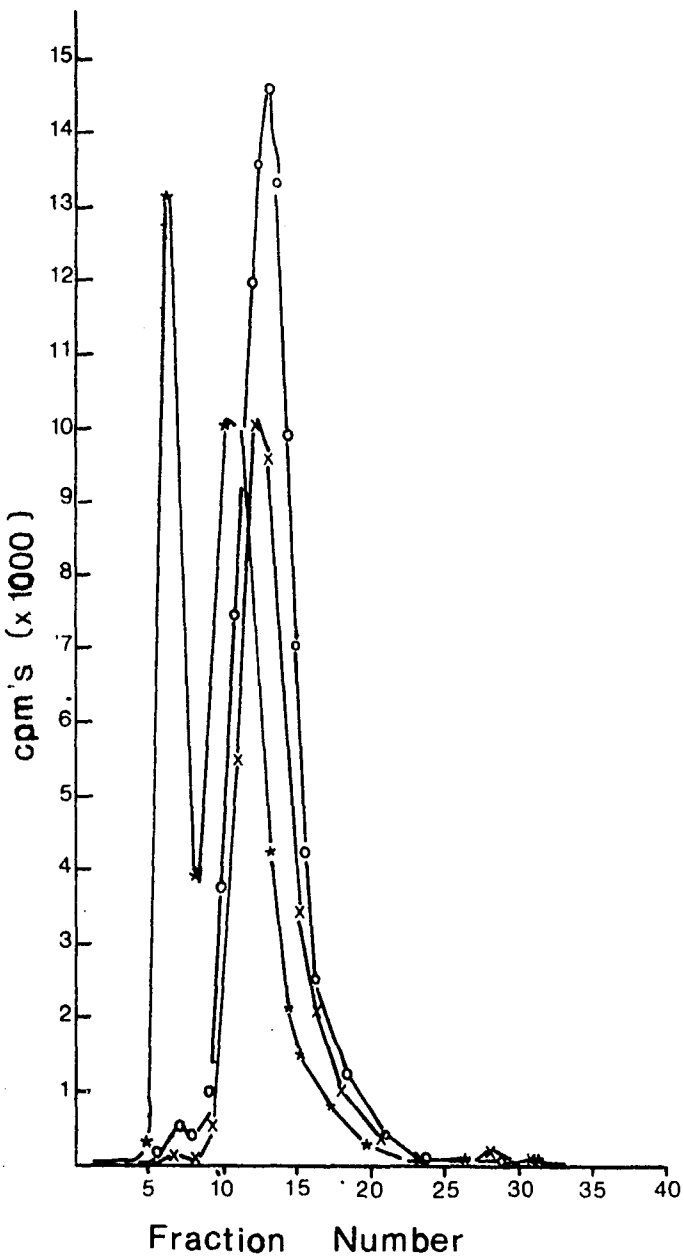
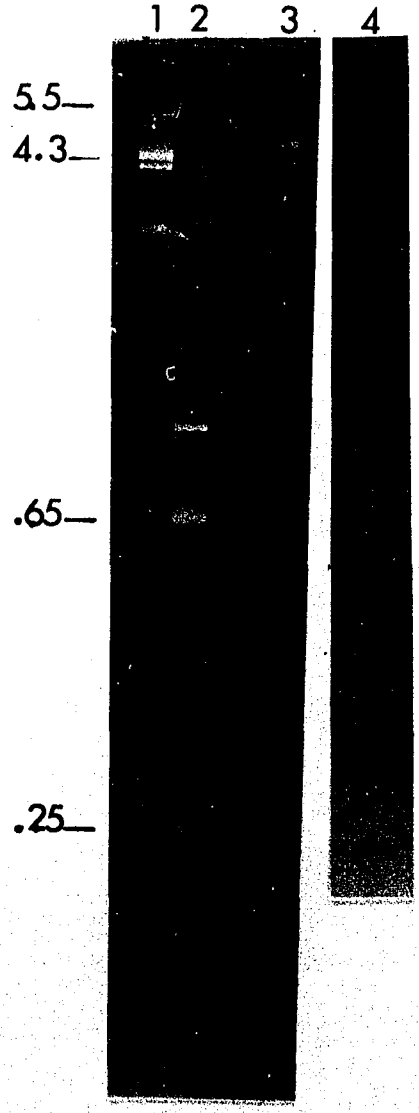

cpm's (x 1000)
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
5
10
15
20
25
30
35
40
Fraction Number

FIGURE 2: Gel and fluorograph of nick translated plasmid DNA using 4.5 units of T4 DNA polymerase I

Lane 1 is plasmid DNA. Lane 2 is an Alu I digest of pBR322 and was used as a length standard. Lengths shown are in kilobases. Lane 3 is nick translated plasmid DNA. Lane 4 is the fluorograph of the DNA in lane 2.


1 2 3 4
5.5
4.3
.65
.25

then treated in one of two ways. Either the sample was immediately added to a 10cm x 1cm Sephadex G-75 column equilibrated in 1x TNE plus 0.1% SDS, or it was stored overnight at -70°C. Before adding to the column, 300 uls of 1x TNE - 0.1% SDS were added to the Eppendorf tube containing the nick translation mixture.

Sephadex G-75 was swelled overnight in 1x TNE at room temperature. A glass wool plug was put in the tip of a 10 ml pipet and the G-75 was packed to a height of 10-12 cms. Before each sample was loaded onto the column, the column was extensively washed with 50-60 mls of 1x TNE - 0.1% SDS.

Fractions were collected using a Gilson microfractionator (model FC-80K) set at 35 drops/tube. This resulted in a fraction volume of approximately 0.75 mls. The column was washed with 1x TNE - 0.1% SDS. The nick translated DNA appeared in the excluded volume which was at about 8.0 mls. The unincorporated nucleotides eluted in the next 20 mls. 1.0 ul of each of the first 15 fractions was counted in minivials containing 4.0 mls of Aquasol. A Beckman scintillation counter (model LS-250) with an isoset adjusted to a wide open window was used to count each sample. The peak fractions of the excluded volume (usually fractions 5-11) were pooled and dialyzed overnight against 1 liter of TE. This dialysis removed most of the salt and some of the SDS. The sample was then frozen, lyopholyzed to dryness, redissolved in 0.5 mls of TE, and dialyzed for at least 6 hours against 1 liter of TE. A 1.0 ul sample was then counted, and the rest of the nick translated DNA stored at 5°C for use in various hybridization reactions. Specific activities of the nick translated DNAs

were usually in the range of 10^7-10^8cpms/ugDNA.

Gel Electrophoresis

Acrylamide

Vertical 130mm x 100mm x 1.5mm acrylamide-TBE gels were used for isolating DNA fragments for cloning, or for analyzing restriction enzyme digests. A 40% acrylamide stock solution was made up in distilled, deionized water. The acrylamide to bisacrylamide ratio was 37.5:1. The following was the procedure used for making one 7% acrylamide 130mm x 100mm x 1.5mm gel:

7 mls of 40% acrylamide stock
4 mls of 10x TBE
29 mls of distilled, deionized water

These were mixed in a 250 ml glass beaker with 0.4 mls of a 10% ammonium persulfate (Biorad) stock solution. Immediately after mixing, 40 uls of TEMED (Biorad) were added. The solution was mixed well and poured between the glass plates. The comb was inserted and the acrylamide allowed to polymerize at least 30 minutes before using the gel.

Usually, a 7-tooth comb (Dankar Plastics) was used. This allowed up to 70 uls of sample and 8-10 ugs DNA per well. For large DNA fragments of 1400-4000 bp, the gel was electrophoresed at 150V for 6-8 hours. For small fragment sizes of 100-1000 bp, the BPB dye mix was made 0.01% xylene cyanol, and the gel was electrophoresed at 150V until the xylene cyanol was 2 inches from the bottom of the gel.

Agarose

Several percentages of agarose gels were used. For large DNA fragments, greater than 4000 bp, or for single stranded viral DNA, gels of 0.5%, 0.7%, or 1% agarose were used. For DNA fragments less than 4000 bp, a 2% agarose gel was used.

Gels were electrophoresed in a horizontal gel apparatus (Dankar Plastics) with a tray that allowed gel lengths of up to 23 cms (Appendix B). Full length gels were formed by adding an appropriate amount of dry agarose to 150 mls of TBE and autoclaving in 250 ml Erlenmeyer flasks. The flasks were autoclaved for 10-15 minutes to thoroughly dissolve the agarose. After autoclaving, the gel solution in the flasks was mixed well by swirling the flask. This was especially important with the higher percentage gels. The hot gel solution was poured into the gel forming tray. The agarose was allowed to solidify at room temperature for 30 minutes to 2 hours before removing the end pieces and the comb.

When the agarose had solidified, the gel tray was placed in the horizontal apparatus and submerged in TBE. BPB dye mix was added to each sample. The samples were loaded into the gel wells and electrophoresed at 30V overnight for 0.5%, 0.7%, and 1% gels and 150-200V for 3-4 hours for the 2% gels.

The gels were stained for 15 minutes in 1 ug/ml ethidium bromide in TBE, and the DNA visualized under UV light. If the DNA bands were not fully resolved, the gels were placed back in the trays and electrophoresed for a longer period of time. Kodak Royal Pan film was used to photograph the gels. Gels could also be wrapped in Saran Wrap and stored in the cold

for several days without dehydration or significant DNA diffusion.

Electroelution of DNA Fragments

DNA bands were sometimes cut out of 7% acrylamide-TBE gels and used for cloning into M13, for nick translations, or for restriction enzyme digests. 8-10 ugs of clone DNA were digested with an appropriate restriction endonuclease in a volume of 40 uls. After digestion, 10 uls of the BPB dye mix were added, and the 50 uls of DNA-dye mixture were loaded into each well. The gel was electrophoresed and stained with ethidium bromide. To avoid too much damage to the DNA, UV illumination was as brief as possible.

The DNA band(s) of interest were cut out of the gel and transferred to a piece of parafilm. These acrylamide pieces were then placed in dialysis tubing (Fisher:Spectropor 0.67inf or 0.25inf). 0.6-0.8 mls of TE and one gel piece was added to each dialysis bag. These dialysis bags were put into an electroelution box with 250-300 mls of 1/10x TBE, and the DNA was eluted from the acrylamide pieces by maintaining a constant current of 35mA (milliamps) overnight. The polarity was then reversed for 1 minute to loosen any DNA stuck to the dialysis bag, and the solution containing the DNA was removed with a Pasteur pipet. The empty bag was rinsed with a small amount of TE and this was combined with the other DNA solution. 20 ugs of tRNA were added as carrier and the DNA ethanol precipitated.

Southern Transfer and Hybridization

Restriction endonuclease digested DNA or viral single stranded DNA was electrophoresed on agarose gels. The DNA was denatured by soaking the gel in 1.5 M NaCl, 0.5 M NaOH for 15-30 minutes with occasional rocking. The alkaline solution was decanted, the gel rinsed with deionized water, and then neutralized by soaking for 30-60 minutes in 2.0 M NaCl, 1.0 M TRIS-HCl (pH5.5).

A transfer apparatus (diagrammed in Appendix C) was set up by filling a pan with 20x SSC. A plastic tray was inverted and placed in the pan containing the SSC. A piece of Whatmann #1 filter paper (Millipore) was soaked in 20x SSC and placed on the inverted plastic tray with the ends of the filter paper hanging in the 20x SSC solution. The filter paper was checked to make sure there were no air bubbles trapped between it and the plastic tray.

The soaked gel was then placed on the filter paper, making sure no air bubbles were trapped. A Kimwipe was used to blot extra liquid from the top of the gel. Then strips of Saran Wrap were placed around the gel covering the filter paper. This prevented the paper towels from touching the wet filter paper.

A piece of nitrocellulose or Genatran 45 was cut 1cm longer and wider than the gel and floated on distilled deionized water. Gloves were worn when cutting and handling the nitrocellulose or Genatran, because fingerprints prevented wetting and DNA transfer. The nitrocellulose or Genatran was laid on top of the gel, again checking for trapped air bubbles. A piece of Whatmann #1 filter paper was cut the same size as the

nitrocellulose and wetted with 20x SSC. This was then placed on top of the nitrocellulose paper. A stack of paper towels at least 5 inches high was placed on top of the Whatmann #1 filter paper. A plexiglas plate was placed on top of the paper towels and this was weighted down with at least a 500 gram weight. Transfer of the DNA from the gel to the nitrocellulose or Genatran paper was allowed to proceed overnight at room temperature.

At the completion of the transfer, the nitrocellulose or Genatran was placed in a pan of 100 mls of 2x SSC and gently rocked for about 2 minutes. The nitrocellulose or Genatran sheet was blotted between paper towels. The sheet was then placed between paper towels and these were placed between two glass plates. The plates were clamped together, and baked in a vacuum oven at 80^{o}C for 2 hours. The baked sheet was soaked in 3x SSC for 30 minutes and then for 3 hours in 1x Denhardts plus 0.2% SDS. Prehybridization using heterologous DNA was not necessary. For Genatran 45, the BSA concentration was raised to 0.5% and the SDS was raised to 1%. After 3 hours, the filters were blotted and air dried.

Prior to hybridization, the filters were wetted in 1x Denhardts and placed in Seal-A-Meal bags with 10-30 mls of 1x Denhardts. The radioactive probe was boiled for 5-10 minutes, cooled on ice, and then added to the bag. The bag was sealed, and the whole mixture was incubated at 60^{o}C overnight.

After hybridization, the filters were washed for 2 hours at 50^{o}C with gentle shaking. Four changes of 2x SSC were made during the 2 hour wash. The filters were then air dried and autoradiographed.

DNA Sequencing

Dideoxy method

An M13mp9 sequencing kit was purchased from BRL and the manual (1980) was followed with several modifications which are described below.

Single stranded, viral DNA template was prepared from 5 mls of JM103 infected culture as described in a previous section. The dried DNA pellet was resuspended in 9.3 uls of sterile, distilled, deionized water in a sterile, 1.5 ml Eppendorf tube. 2 uls of the 16bp universal primer were added along with 1.2 uls of the 10x concentrated Hin buffer. The contents were mixed well by tapping the tube with a pencil. The mixture was concentrated at the bottom of the Eppendorf tube by a brief spin in the microfuge. The sample was then heated for 10 minutes in a 85-95^{o}C water-filled heating block. After 10 minutes, the whole block was removed, placed on the bench top, and allowed to cool slowly to 30^{o}C. This allowed the annealing of the primer to the template and usually took 45-60 minutes. The liquid in the tubes was again spun down for 30 seconds in the microfuge.

While the samples were cooling in the heating block, 4 microfuge tubes (labelled G, A, T, C) were prepared by adding the following reagents to each tube:

Tube G: 1 ul G' and 1 ul ddG
Tube A: 1 ul A' and 1 ul ddA
Tube T: 1 ul T' and 1 ul ddT
Tube C: 1 ul C' and 1 ul ddC

These solutions were added to opposite sides of the tubes to aviod mixing. As noted in Appendix B, it was very important to make fresh

nucleotide stock solutions every two weeks.

After the annealing sample had cooled, 1-2 uls (5-10uCi) of [^{32}P]dATP (Ammersham: > 800Ci/mmol) were added to the annealed mixture along with 1 ul of 0.1 M dithiothreitol and 1 ul of DNA polymerase Klenow fragment. The contents were mixed by pipetting. 3 ul aliquots were then put into each of the 4 tubes containing the deoxy and dideoxy mixes. The 3 uls were put in the tube so that it would not mix with either the deoxy or dideoxy mix. The tubes were then centrifuged in the microfuge to simultaneously mix the annealed, deoxy, and dideoxy solutions in all four tubes. The DNA sequencing reactions were then allowed to proceed for 30 minutes in a 30^{o}C water bath.

After this incubation, 1 ul of 0.5 mM dATP was added to each of the four tubes, the tubes centrifuged, and then incubated for another 30 minutes in the 30^{o}C water bath. This extended the DNA chains which were prematurely terminated due to low dATP concentrations.

The sequencing reactions were stopped by addition of 14 uls of formamide-dye mix which was made 20 mM NaOH just prior to use. The DNA in the samples was then denatured by heating at 95-100^{o}C in a water filled heating block. After denaturing for 5 minutes, 4-5 uls of each sample were loaded into separate gel wells using a 10 ul Hamilton syringe. The rest of the sample was stored at -20^{o}C. The syringe was extensively rinsed after loading each sample. The samples were loaded in consecutive wells in the order G, A, T, C. This was electrophoresed at 1600V, 40mA, for 2 to 2-1/2 hours, or until the bromophenol blue reached the bottom of the gel.

Maxam-Gilbert method

The Maxam-Gilbert DNA sequencing was performed in Dr. John Donelson's laboratory at the University of Iowa, Iowa City, Iowa.

End labelling Two fragments of pPGA7 were purified by electroelution. The 0.25 kb Hpa II-Hpa II fragment, and the 0.42 kb Bam HI-Hinc II fragment were end labelled with [^{32}P]dCTP or [^{32}P]dATP.

10 ugs of 0.25 kb Hpa II-Hpa II fragment were first digested with 50 units of Hae III in a total volume of 100 uls. The digest was allowed to proceed for 1 hour at 37^{o}C. This was then heat inactivated at 65^{o}C for 10 minutes. This digest produced two fragments; one 0.19 Kb and the other 0.06 kb. The following amounts of the following solutions were then combined:

100 uls Hae III digested 0.25 kb Hpa II-Hpa II fragment
10 uls 10x 'L' buffer
6 uls 5.0 M NaCl
78 uls sterile distilled water
4 uls [^{32}P]dCTP (New England Nuclear: > 3000Ci/mmol)
2 uls DNA polymerase I (1 unit/ul)

The total reaction volume was 200 uls. This was mixed and incubated on ice for 10 minutes. 2 uls of 1.0 mM dCTP were then added and the solution mixed and incubated on ice for another 5 minutes. The recognition sequence for Hpa II is 5'-CCGG-3' with the enzyme cutting after the first C. This leaves a GC exposed on the 3'-5' DNA strand. By adding radioactive dCTP in the above reaction, the 5'-3' strand was labelled.

The 0.42 kb Bam HI-Hinc II fragment was labelled in the following manner:

Precipitated DNA fragment (10 ugs)
20 uls 10x 'XH' buffer
172 uls sterile distilled water
4 uls [^{32}P]dATP (New England Nuclear: > 3000Ci/mmol)
2 uls 10 mM dGTP
2 uls DNA polymerase I (1 unit/ul)

This was mixed and incubated on ice for 15 minutes. The recognition sequence for Bam HI is 5'-GGATCC-3', and that for Hinc II is 5'-GTPyPuAC-3'. Bam HI cuts after the first G, leaving 'sticky' ends, while Hinc II cuts between the pyrimidine and purine, leaving 'blunt' ends. In the above labelling reactions, the Bam HI site was preferentially labelled. The Bam HI digestion leaves single stranded DNA sequence of CTAG on the 3'-5' strand. The unlabelled dGTP base paired with the C and then the polymerase added the radioactive dATP to base pair with the T on the 3'-5' strand.

The labelling reactions for both fragments were stopped by addition of 200 uls of sterile distilled water. Each reaction was then phenol extracted two times followed by two ether extractions. 1.0 ml of ice cold 100% ethanol was added and the DNA fragments were precipitated. The precipitated, labelled DNA fragments were then resuspended in 100 uls of sterile, distilled water. 10 uls of 5.0 M ammonium acetate (pH8) and 250 uls of ice cold 95% ethanol were then added to each resuspended fragment, and the DNA precipitated again. Each DNA pellet was again resuspended in 100 uls of sterile, distilled water, 10 uls of 5.0 M ammonium acetate (pH8), and 250 uls of 95% ethanol, and the DNA precipitated again. After these precipitations, the specific activity of the Hpa II-Hpa II fragment was 1.01×10^6cpm/ugDNA, and that of the Bam HI-Hinc II fragment was 2.26×10^6cpm/ugDNA.

The Hpa II-Hpa II fragment was resuspended in 120 uls, and the Bam HI-Hinc II in 360 uls of sterile water. This was divided among four Eppendorf tubes labelled G, C, A, and T. (The G tube was for the G reaction; the C tube was for the C reaction; the A tube was for the A+G reaction; and the T tube was for the T+C reaction.) 20 uls of the labelled fragment were put into each of the tubes labelled G and C. 40 uls were put into each of the tubes labelled A and T. The extra Bam HI-Hinc II solution was stored at -70°C.

DNA sequencing 1.0 ul of 1 mg/ml tRNA in water was added to each of the four reaction tubes. Three volumes of an ice cold 90% ethanol, 100 mM $MgCl_2$, 0.3 mM ammonium acetate (pH5.5) solution were added to each tube and the tubes put in dry ice for 10 minutes to precipitate the DNA. The tubes were then centrifuged in the microfuge in the cold for 15 minutes, and the supernatant removed with a drawn out Pasteur pipet. 0.5 mls of ice cold 70% ethanol were added, the samples centrifuged for 5 minutes, the supernatant removed, and the pellets dried in a vacuum desiccator for 10 minutes. The pellets in the 'T' tubes were resuspended in 20 uls water. 60 uls of ice cold 95% ethanol were added, the samples put in dry ice for 10 minutes, and then centrifuged and dried as before.

The following base modifications were then carried out:

G tube: Resuspend DNA pellet in 200 uls G-mix
Add 1 ul ice cold DMS
Incubate at 20°C for 4 minutes then put on ice
Immediately add 50 uls G-stop, 4 uls BME, vortex
Add 750 uls cold 95% ethanol, vortex, dry iced

C tube: Resuspend DNA pellet in 20 uls C-mix
Add 30 uls cold hydrazine, mix
Incubate at 20°C for 8 minutes then put on ice
Immediately add 200 uls of CT-stop, vortex
Add 750 uls cold 95% ethanol, vortex, dry iced

T tube: Resuspend DNA pellet in 20 uls T-mix
Add 30 uls cold hydrazine, mix
Incubate at 20°C for 4.5 minutes, then put on ice
Immediately add 200 uls of CT-stop, vortex
Add 750 uls cold 95% ethanol, vortex, dry iced

A tube: Resuspend DNA pellet in 100 uls A-mix
Add 1 ul of 1 mg/ml sheared, denatured salmon sperm DNA
Seal tube with teflon tape, place in press, tighten
Put press in 90°C water bath for 3.5 minutes
Plunge press into ice.
Add 150 uls A-stop, 5 uls 1 mg/ml tRNA, vortex
Add 750 uls cold 95% ethanol, vortex, dry iced

All samples were left on dry ice for at least 15 minutes. They were then centrifuged for 15 minutes, and the supernatants transferred to separate Eppendorf tubes. The counts in the DNA pellets were checked before discarding the supernatants. At this time, there was about a 90% recovery of the starting counts. The pellets were then resuspended in 200 uls of 0.3 M sodium acetate (pH5.5). 600 uls of ice cold 95% ethanol were added, the samples vortexed, and the DNA precipitated by putting the samples in dry ice for 10 minutes. The samples were then centrifuged for 15 minutes, the supernatant removed, the radioactivity in the pellets checked, and the pellets resuspended in 300 uls of 0.3 M sodium acetate (pH5.5). 300 uls of ice cold 95% ethanol were added to each tube, the tubes vortexed, and the DNA precipitated as before. After the radioactivity in the pellets was checked, the pellets were dried in a vacuum desiccator for 30 minutes.

50 uls of 1.0 M piperidine were added to each tube containing the

dried DNA pellets. The tubes were sealed with teflon tape, vortexed and put into the press. The press was placed in a 90°C water bath for 30 minutes. At the end of the incubation, the press was plunged into ice. The tubes were removed from the press and given a quick spin in the microfuge to bring down the solution condensed at the top of the tube. The piperidine solution was removed and put into new Eppendorf tubes. The Eppendorf tips used to transfer the piperidine were saved. 50 uls of 0.3 M sodium acetate (pH5.5) were added to the 'old' tubes. The tubes were vortexed, and the previously used Eppendorf tips were used to transfer the sodium acetate solution to the 'new' tubes containing the piperidine solution. The radioactivity in the 'old' and 'new' tubes was checked. If there were too many cpms in the 'old' tubes, 25 uls of the sodium acetate solution were added, the tubes vortexed and the solution added to the 'new' tubes. The Eppendorf tips and the 'old' tubes were then discarded. Each solution in the 'new' tubes was vortexed. 400 uls of ice cold 95% ethanol were then added to each tube, the tubes vortexed, placed in dry ice for 10 minutes and centrifuged for 15 minutes. The supernatant was removed and the pellets dried in a vacuum desiccator for at least 1 hour.

The cpms in each DNA pellet were checked using a hand held Geiger counter. The DNA pellets were resuspended in appropriate volumes of MG dye mix so that the cpms in the A and T tubes were twice that of the cpms in the G and C tubes. The resuspended samples were then put into the press and put in boiling water for 1 minute. The press was plunged into ice and 1 ul of each sample was loaded into gel wells. The samples were loaded in consecutive wells in the order: A, G, C, T.

DNA Sequencing Gels

The DNA sequencing gels for the Maxam-Gilbert sequencing procedure were 85 cm long. They were prepared by the technicians in Dr. John Donelson's lab.

The DNA sequencing gels for the M13 dideoxy procedure were 430mm x 170mm x 0.5mm. The glass plates were purchased from Central Stores at Iowa State University, and the spacers and combs were purchased from DanKar Plastics.

The glass plates for the dideoxy sequencing gels were washed carefully. One side of each plate was scrubbed with a paste of Comet. The plates were then rinsed extensively with distilled water and Kimwipes to remove all the Comet. Each plate was dried with Kimwipes and then 95% ethanol was squirted on the cleaned glass. Kimwipes were used to spread the ethanol on the glass. The plates were wiped with clean Kimwipes until the ethanol had evaporated. The cleaned spacers were then placed along the sides and the bottom of the longer plate. The shorter plate was placed cleaned-side down on the long plate, the spacers adjusted, and both plates clamped together.

To pour one dideoxy sequencing gel, 8 mls of a 40% acrylamide stock solution (19:1 acrylamide to bisacrylamide) were mixed with 10 mls distilled, deionized water, 4 mls of 10x TBE, and 19 grams of urea. This solution was put into a 250 ml glass beaker, warmed in a 42^{o}C water bath to dissolve the urea, and then filtered through three layers of Miracloth into a graduated cylinder. Distilled, deionized water was then added to the cylinder to give a final volume of 40 mls. The solution was poured

back into the glass beaker and mixed well by swirling. 0.2 mls of a 10% ammonium persulfate stock solution in water were added to the beaker with the acrylamide-urea solution. This was mixed well, and then 20 uls of TEMED were added and the gel immediately poured. The comb was inserted and the acrylamide allowed to polymerize at room temperature for 1 hour in a horizontal position with the comb end resting on a petri dish. If gels were poured a day ahead of time, Saran Wrap was placed over the comb end and the gel stored overnight in the cold. When gels were prepared ahead of time, they were allowed to warm up to room temperature before removing the comb.

Just before use, the combs and bottom spacer were carefully removed and the wells immediately flushed with 1x TBE in a 30 ml syringe with a 20-gauge needle. Silicone vacuum grease was smeared on the gel apparatus (Appendix B) and the gel clamped on. The short glass plate was placed toward the vacuum grease. Clamps were placed very close together along both vertical sides of the gel. If two gels were being run simultaneously, then the other gel was placed on the other side of the apparatus in the same manner, otherwise a glass plate was clamped to the other side. The clamping of the gels to the apparatus formed the upper buffer chamber. This chamber was then filled with 1x TBE and the apparatus checked for leaks. The apparatus was placed in the bottom buffer reservoir which also contained 1x TBE. The negative electrode was attached to the upper reservoir and the positive electrode to the bottom reservoir. The protective shield was placed around the apparatus, and the gels were pre-electrophoresed (using a Dankar power supply provided by Dr.

Jack Horowitz, Iowa State University) at 1600V or 40-50mA until the temperature of the glass plates reached 50-60°C. This usually took 2-3 hours.

Just prior to loading the samples, the current was turned off and the wells were flushed with 1x TBE. To prevent 'smiling' of DNA bands, samples were not loaded into the four outermost wells. These, and any other blank wells, were filled with 4-5 uls of the formamide-dye mix. After loading all the sequencing samples, the gels were electrophoresed at 1600V or 40mA for 2 or/and 5 hours depending on the length of DNA sequence to be read.

At the end of the run, the upper buffer was poured out and the clamps removed. The long gel plate was cooled under running tap water for a few minutes. This usually caused the gel to adhere to the long plate. The glass plates containing the gel were placed on newspaper, long-plate-side down, and a metal spatula was used to pry the plates apart. This was done carefully and slowly to prevent the gel from tearing or air bubbles from forming between the gel and the long glass plate. If air bubbles did form, they were carefully worked out using a wash bottle of TBE and metal spatula.

A piece of Whatmann #1 filter paper was placed on the gel. The paper was rubbed firmly over the gel which caused the gel to stick to the filter paper. The filter paper plus gel was carefully peeled away from the long glass plate. At this point, a Geiger counter could be used to check the amount of radioactivity present. The gel was then dried for 1 hour using a large Hoeffer gel drier supplied by Dr. Joan Stadler (Iowa State

University). The dried gel was autoradiographed at room temperature overnight. Sometimes the gel was a bit sticky after drying, so a sheet of the paper packed with the sheets of X-ray film was placed between the gel and the film.

Determination of the Repeated Sequence Size in pPGA7 and pPGB9

To determine the size of the repeated sequence contained in clones pPGA7 and pPGB9, trace amounts of nick translated Bam HI inserts of pPGA7 and pPGB9 were hybridized to unlabelled calf thymus DNA. The reactions were performed in 100 ul volumes. 90 uls of 3.0 A_{260} units of calf thymus DNA in TE were mixed with 10^6cpm of clone insert. (Control hybridizations of 10^6cpm of clone insert alone were run in parallel.) The DNA in the reactions was denatured by placing the tubes in boiling water for 10 minutes. The tubes were immediately put on ice, and 10 uls of 3.0 M NaCl were added to each tube. The tubes were incubated at 60°C for 2 hours.

Two tubes each of reactions containing calf thymus plus pPGA7 insert, calf thymus plus pPGB9 insert, pPGA7 insert alone, and pPGB9 insert alone were not incubated at 60°C. These served as 'initial time' or 'zero time' controls.

20 ugs of a 4 mgs/ml stock of tRNA were added to each reaction tube either immediately after boiling (control tubes) or after the hybridizations. All samples were then ethanol precipitated.

The dried DNA pellets from each hybridization and control sample were resuspended in 10 uls of sterile water. 2 uls of 10x S1 nuclease assay

buffer, 0 or 70 units of S1 nuclease, and sterile water were then added to each tube to give a final volume of 20 uls. All samples were incubated at 37^{o}C for 2 or 4 hours. 5 uls of BPB dye mix were added to each sample, and the samples electrophoresed on a 2% agarose-TBE gel. The gel was stained with ethidium bromide, wrapped in Saran Wrap and autoradiographed at room temperature.

Hybridization Experiments to Determine the Fraction of the Bovine Genome Complimentary to pPGA7 and pPGB9

Viral DNA from M13 clones containing the Bam HI insert of pPGA7 or pPGB9 was hybridized to trace amounts of nick translated calf thymus DNA. Hybridization reactions were performed in 100 ul volumes in 0.3 M NaCl. Samples containing viral DNA plus 10^{6}cpms of calf thymus DNA, and samples containing just 10^{6}cpms of calf thymus DNA were boiled for 10 minutes, put on ice, then either ethanol precipitated or incubated at 60^{o}C for lengths of time corresponding to Cot values of 10 and 100. (The Cot $_{1/2}$ for repetitive bovine DNA is 0.1.)

The concentration of the viral DNA stock containing the Bam HI insert of pPGA7 (called M13mp8BA7) was 4.1 A_{260} units and its length was about 9 Kb. The concentration of the viral DNA stock containing the Bam HI insert of pPGB9 (called M13mp8BB9) was 2 A_{260} units and its length was about 11 Kb. The calculated Cot $_{1/2}$ values for the M13mp8BA7 and M13mp8BB9 were 2.1×10^{-3}, and 2.6×10^{-3}, respectively.

Three sets of hybridization experiments were set up as follows:

	Sample	Hybridization time	Equivalent Cot
#1	M13mp8BA7 plus 10^8 cpms of calf thymus DNA	2 hours 6 minutes	10
		21 hours 4 minutes	100
#2	M13mp8BB9 plus 10^8 cpms of calf thymus DNA	4 hours 19 minutes	10
		43 hours 12 minutes	100
#3	M13mp8BA7 plus M13mp8BB9 plus 10^8 cpms of calf thymus DNA	1 hour 26 minutes	10
		14 hours 24 minutes	100

Controls of 10^6 cpms of calf thymus DNA alone were incubated at $60^{\circ}C$ along with each of the above experiments. At the end of the hybridization incubation, all samples were ethanol precipitated. The dried DNA pellets were incubated at $37^{\circ}C$ with 0 or 70 units of S1 nuclease in a total reaction volume of 100 uls. 100 uls of cold 0.1 M potassium phosphate buffer (pH6.8) were added to each sample, the sample applied to a hydroxylapatite column and the amount of radioactivity in the hybridized DNA determined.

Since the M13 viral DNA was used, self-hybridization of the viral DNA could not occur. Only 1 strand of the cloned bovine repeated sequence was available for hybridization; therefore, values obtained for the percent of bovine DNA which hybridized were doubled to give the percent of the bovine genome made up of repeated sequences.

Hydroxylapatite Chromatography

An apparatus was built whereby 18 hydroxylapatite (HAP) columns could be run at the same time. This apparatus used plastic disposable syringes and is shown in Appendix C.

Each column used 2.0 grams of DNA grade HAP (Biorad) which had been boiled for 15 minutes in 0.2 M potassium phosphate buffer (pH6.8) (Kohne and Britten, 1975; Britten, et al., 1974). After boiling, the HAP was equilibrated with 0.05 M potassium phosphate buffer (KPB), and the fines removed before pouring the column. The 2.0 grams of equilibrated HAP was then poured into the syringe and the HAP allowed to settle. Each column was washed twice with 2.0 mls of 0.05 M KPB. The second wash was collected in a plastic minivial, and this was used as a control.

The DNA was then added to the column in 0.05 M KPB (40 ugs of tRNA were added to those samples which contained less than 50 ugs of DNA). The column was washed with 2.0 mls of 0.05 M KPB at room temperature, then with 2.0 mls of 0.05 M KPB at 60^{o}C, and then twice with 2.0 mls of 0.2 M KPB at 60^{o}C.

3.0 mls of Aquasol scintillation cocktail were then added to each of the following minivials:

#1: 2.0 mls of 0.05 M KPB wash before addition of the DNA sample

#2: 2.0 mls of 0.05 M KPB at room temperature after the addition of the DNA sample

#3: 2.0 mls of 0.05 M KPB at 60^{o}C after the addition of the DNA sample

#4: 2.0 mls of 0.2 M KPB at 60^{o}C after the addition of the DNA sample

#5: 2.0 mls of 0.2 M KPB at 60°C after the addition of the DNA sample

After the cocktail was added to each of the vials, the solution in the vials was mixed well. These were placed in a scintillation counter and each vial counted for 5-10 minutes. The percentage of double stranded DNA in each hybridization reaction was then determined by dividing the cpms in vials #4 + #5 by the total cpms in the sample, vials #2 + #3 + #4 + #5.

S1 Nuclease Digestion

S1 nuclease was obtained from BRL. Under appropriate conditions of salt and temperature, S1 nuclease selectively degrades single stranded but not double stranded DNA. Figure 3 demonstrates that the conditions used in this work were selective for single stranded DNA.

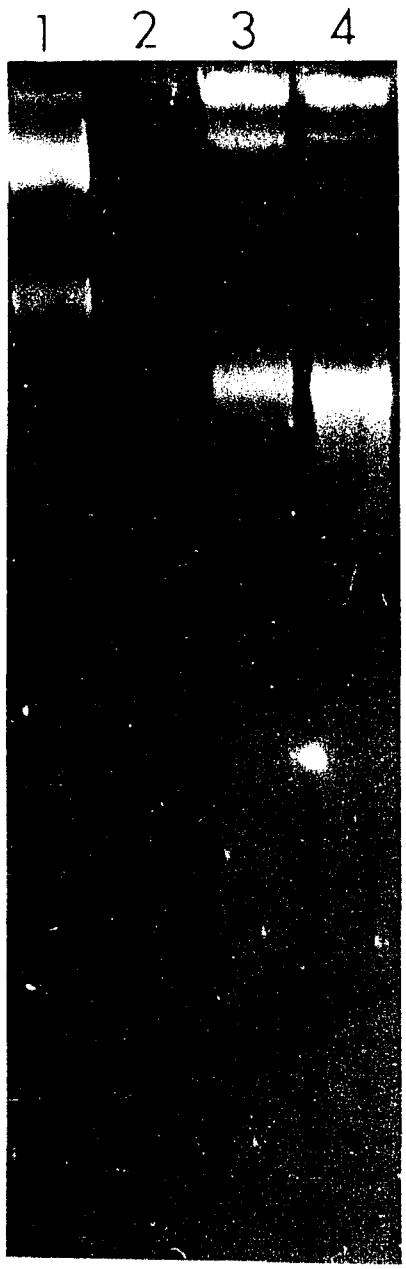
For digests of single stranded DNAs, 1/10 volume of the 10x S1 nuclease assay buffer, and at least 0.5 units of enzyme/ug of DNA were used per reaction. The digestion mixtures were incubated at 37°C for 1, 2, or 4 hours. The reactions were stopped by ethanol precipitation for gel electrophoresis or dilution with KPB for hydroxylapatite chromatography.

Ethanol Precipitation of DNA

One-tenth volume of 3 M sodium acetate (pH8) was added to the DNA solution to be precipitated. At least two volumes of ice cold 100% ethanol were then added, and the tubes containing the DNA, sodium acetate,

FIGURE 3: Double and single stranded DNA digested with S1 nuclease under the conditions used in all S1 nuclease digestions

Lane 1, single stranded, M13 viral DNA; Lane 2, single stranded, M13 viral DNA plus S1 nuclease; Lane 3, double stranded, M13 replicative form DNA; Lane 4, double stranded, M13 replicative form DNA plus S1 nuclease.


1
2
3
4

and ethanol were mixed well. These were then placed at -70°C for at least 30 minutes. This was followed by a 15-20 minute centrifugation in the microfuge or the IEC centrifuge at 8,000-10,000rpms. The supernatant was poured off, and 1.0 ml of ice cold 70% ethanol was then added without mixing. This was centrifuged for another 5 minutes, the supernatant poured off, the top of the tube covered with a Kimwipe, and the pellets dried under vacuum for 5 minutes.

If DNA concentrations in the initial solution were low, or if small pieces of DNA (< 500 bp) were to be precipitated, 20-40 ugs of a 4 mg/ml tRNA solution in water were added to act as a carrier.

Spermine Precipitation of DNA

Fragments of DNA greater than 300 bp were effectively precipitated using spermine. The salt concentration of the DNA to be precipitated was adjusted to 50 mM. Appropriate volumes of a 0.1 M stock solution of spermine in water were added to give a final concentration of 5 mM. This solution was mixed, incubated on ice for 15 minutes, and then centrifuged in the microfuge for 15 minutes.

The supernatant was poured off and 1.0 ml of spermine wash was added to the pellet. This was incubated on ice for at least 2 hours. The DNA-wash solution was centrifuged for 15 minutes in the microfuge and the supernatant again poured off. 1.0 ml of ice cold 70% ethanol was added to the pellet, and this was again centrifuged for another 5 minutes. The supernatant was poured off, and the DNA pellet dried.

Fluorography

When tritiated DNA was used as a hybridization probe, the nitrocellulose filter paper was impregnated with a fluor. This was necessary in order for the tritium beta emissions to expose the X-ray film.

After the hybridization, the nitrocellulose paper was thoroughly air dried. The dried sheet was then saturated with a 30% solution of 2,5-diphenyloxazole (PPO) in toluene. The fluor saturated nitrocellulose was air dried and then taped to a piece of cardboard. A piece of X-ray film was placed against the nitrocellulose and placed between two pieces of cardboard. This was wrapped in aluminum foil, and the light-tight package was clamped and kept at -70^{o}C for 3-4 weeks. After the exposure time, the X-ray film was developed.

Photography

Photographs of agarose and acrylamide gels were taken with a Polaroid camera positioned over a long wave UV light box (UV Products). A red filter was made from red cellophane and used in taking the pictures.

Two types of film were used, Polaroid type 57 or Kodak Royal Pan 4141 (Estar thick base, 0.18mm). Type 57 film was exposed from 1/2 to 3 seconds depending upon the brightness of the DNA bands. It was found that the Kodak film gave higher quality photographs, so it was used for all later photographs.

The Kodak Royal Pan negatives were loaded in total darkness placing

the two notches in the upper right hand corner of the cassette. The film was exposed for 5-30 seconds depending upon the brightness of the DNA bands. The exposed film was then developed in total darkness. The negatives were placed in Kodak D-76 developer for 8 minutes, a stop bath of 3.5% acetic acid for 30 seconds, and fixed in Kodak Fixer for 5 minutes. After fixing, the negatives were rinsed in running distilled water for 30 minutes and air dried.

Selected negatives were printed on Kodak Kodabrome II RC paper.

Purification of Acrylamide

Acrylamide was purified using an ion exchange resin (Biorad, AG501-X8D). For every 100 mls of 40% acrylamide in water, 2-3 grams of resin were added. This was then mixed on a stir plate for 4 hours. The acrylamide-resin mixture was filtered 4 times through 4 changes of Whatmann #1 filter paper. This extensive filtration effectively removed all resin particles. The appropriate amount of bisacrylamide was then added to the acrylamide solution. This acrylamide-bisacrylamide stock solution was stored at $5^{o}C$ in a brown or foil covered bottle. This purification worked well for cleaning up acrylamide used in electroelution gels or DNA sequencing gels.

RESULTS

Restriction Endonuclease Site Mapping of pPGA7 and pPGB9

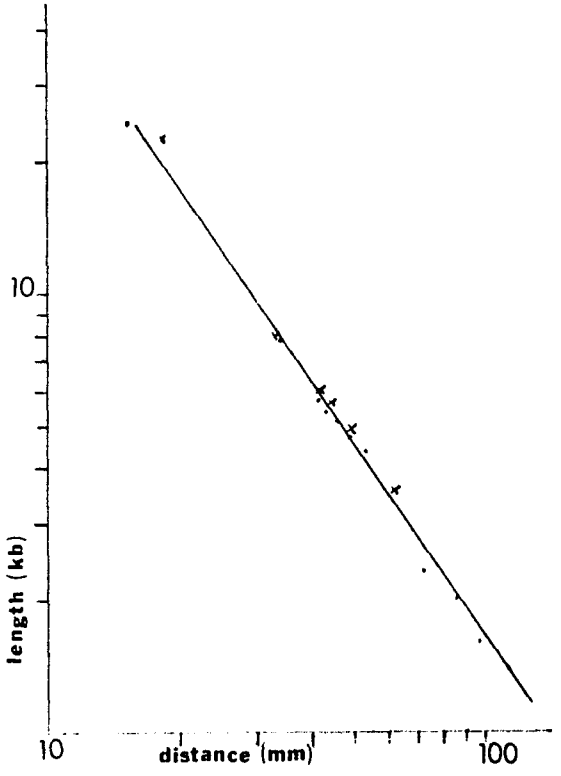
Clones pPGA7 and pPGB9 were digested with various restriction endonucleases, and the location of the restriction sites determined. The sizes of the bovine inserts for pPGA7 and pPGB9 were determined to be approximately 1.4 kb and 3.4 kb respectively (Figure 4).

pPGA7 was first digested with a series of restriction endonucleases which cut pBR322 only once or twice at known locations (Sutcliff, 1978). A sample of the a digest is shown in Figure 5. The results of this and similar digests showed that there were one Eco RI and one Hinc II site, and no Sal I, Hind III, Bgl II, Kpn I, Pst I, and Xba I sites in the bovine DNA insert of pPGA7. Table 1 shows a list of restriction endonucleases used to digest pPGA7 and whether or not they cut pBR322 and/or the bovine DNA insert. Double digestion with Bam HI unambiguously established the location of the Eco RI and Hinc II sites shown in Figure 6. This map was confirmed by electroeluting the Bam HI insert of pPGA7, and digesting this fragment with Hinc II, Eco RI, or Hinc II plus Eco RI. The results of this experiment are shown in Figure 7.

Restriction enzymes such as Hae III and Hpa II cut pBR322 at more than 20 sites (Sutcliffe, 1978), so it is difficult to map the Hae III or Hpa II sites in the bovine DNA insert by digestion of the total plasmid DNA. The Hae III and Hpa II sites were mapped by electroelution of the Bam HI insert, the Hinc II fragments of the insert, and the Eco RI fragments of the insert. These fragments were then digested with

FIGURE 4: Length determination of the bovine DNA insert in clones pPGA7 and pPGB9

pPGA7 (lane 3) and pPGB9 (lane 4) digested with Bam HI and electrophoresed on a 0.8% agarose-TBE gel. A Hind III plus Eco RI digest (lane 1) and an Eco RI digest (lane 2) of lambda DNA were used as length standards. Lengths shown are in kilobases. Also shown is a plot of the standard lengths as a function of the distance migrated from the top of the gel. Standard lengths from lane 1 (.) and from lane 2 (x). This determines the length of the pPGA7 insert as 1.4 Kb, and the length of the pPGB9 insert as 3.4 Kb.


10
length (kb)
10
distance (mm)
100

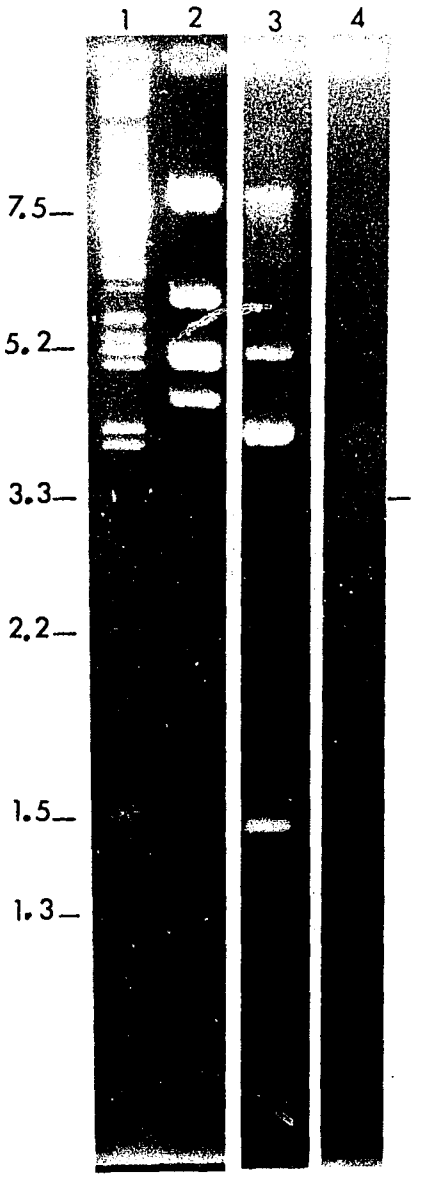

1
2
3
4
7.5
5.2
3.3
2.2
1.5
1.3

FIGURE 5: Restriction endonuclease digests of pPGA7

pPGA7 was digested with various restriction endonucleases and electrophoresed on a 1% agarose-TBE gel. Single (lanes 3-6) and double (lanes 1 and 2) restriction endonuclease digests of pPGA7. The Hind III plus Eco RI digest of lambda DNA (lane 7) is used as a length standard. Lengths in kilobases. Lane 1, Bam HI plus Eco RI; Lane 2, Bam HI plus Hinc II; Lane 3, Hinc II; Lane 4, Hind III; Lane 5, Eco RI; Lane 6, Sal I.

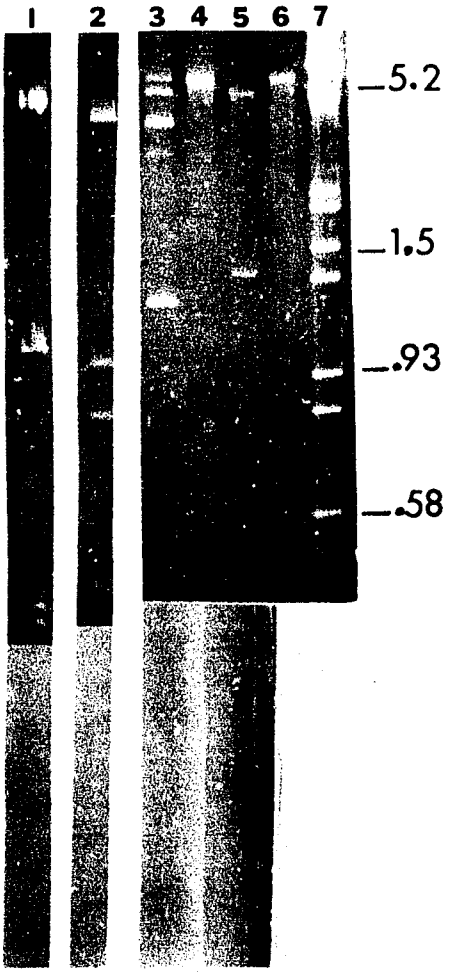

1
2
3
4
5
6
7
5.2
1.5
.93
.58

TABLE 1: Restriction endonucleases which cut (+) or do not cut (-) the pBR322 DNA or the pPGA7 bovine DNA insert.

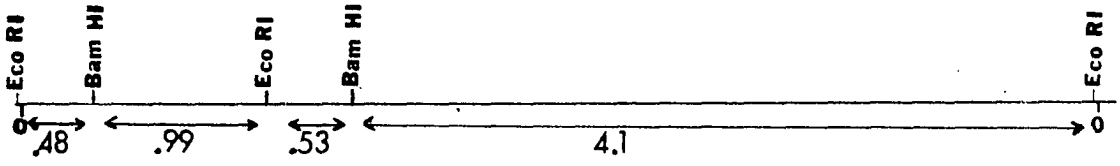
Enzyme	pBR322	pPGA7 Insert
Alu I	+	+
Ava I	+	+
Ava II	+	+
Bam HI	+	+
Bgl II	+	-
Eco RI	+	+
Hae III	+	+
Hinc II	+	+
Hind III	+	-
Hpa II	+	+
Kpn I	-	-
Pst I	+	-
Sal I	+	-
Sau 3A	+	+
Sst II	-	-
Xba II	-	-

FIGURE 6: Restriction endonuclease maps of pPGA7

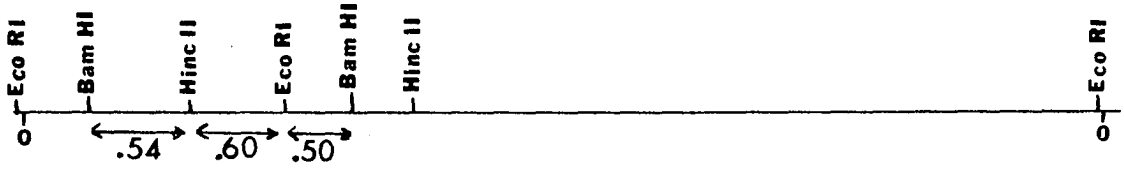
(a) Location of the Eco RI site and fragment sizes (in kilobases) produced by Bam HI Eco RI digest of pPGA7.

(b) Map of the Hinc II and Eco RI sites showing fragment sizes (in kilobases) resulting from a Hinc II plus Eco RI digest of the Bam HI insert of pPGA7.

a


Eco RI
Bam HI
Eco RI
Bam HI
Eco RI
0
.48
.99
.53
4.1
0

b


Eco RI
Bam HI
Hinc II
Eco RI
Bam HI
Hinc II
Eco RI
0
.54
.60
.50
0

combinations of Hae III and Hpa II. Figures 7, 8, and 9 show the digests that were used for this mapping. Hpa II produces fragments of 1.1 kb, 0.25 kb, and 0.06 kb in length, and Hae III yields fragments of 0.6 kb, 0.35 kb, 0.31 kb, and 0.17 kb in length. Table 2 shows the data from Figures 7, 8, and 9. Table 2 also lists the conclusions reached from these data and the subsequent logic used in locating the Hae III and Hpa II sites of the Bam HI insert of pPGA7 (Figure 10).

A restriction map of pPGB9 was determined by Peter Good (graduate student, University of Wisconsin). To map the Pst I sites, pPGB9 was digested with Pst I (Figure 11). This digest is a partial digest, but it suggests that there are three Pst I sites in the bovine DNA insert. The interpretation of the electrophoretic gel in Figure 11 is that the 2.45 kb fragment is generated from the previously mapped Pst I site in the DNA insert and the Pst I site at position 1500 in pBR322. The small fragment of 0.24 kb must represent the distance between two closely spaced Pst I sites in the insert. The faint 0.33 kb fragment can only be explained as a partial digest product of yet a third closely spaced (0.09 kb) Pst I site. The faint 2.8 kb fragment is apparently a partial digest product between the pBR322 Pst I site and the middle Pst I site in the insert. Figure 12 shows the probable locations for these sites derived from the data in Figure 9.

Location of the Repetitious Sequences in pPGA7 and pPGB9

The repetitious sequences in clones pPGA7 and pPGB9 were located by hybridizing Southern blots of electrophoretic gels to ^{32}P-nick translated,

FIGURE 7: Bam HI insert of pPGA7 digested with various restriction endonucleases electrophoresed on a 2% agarose-TBE gel

Lane 1 is an Alu I digest of pBR322, and Lane 2 is a Hae III digest of M13mp8. These were used as length standards. Lengths shown are in Kilobases. Lane 3, Bam HI insert of pPGA7; Lane 4, Hinc II; Lane 5, Eco RI; Lane 6, Hinc II plus Eco RI; Lane 7, Hinc II plus Hae III; Lane 8, Eco RI plus Hae III; Lane 9, Eco RI plus Hpa II.

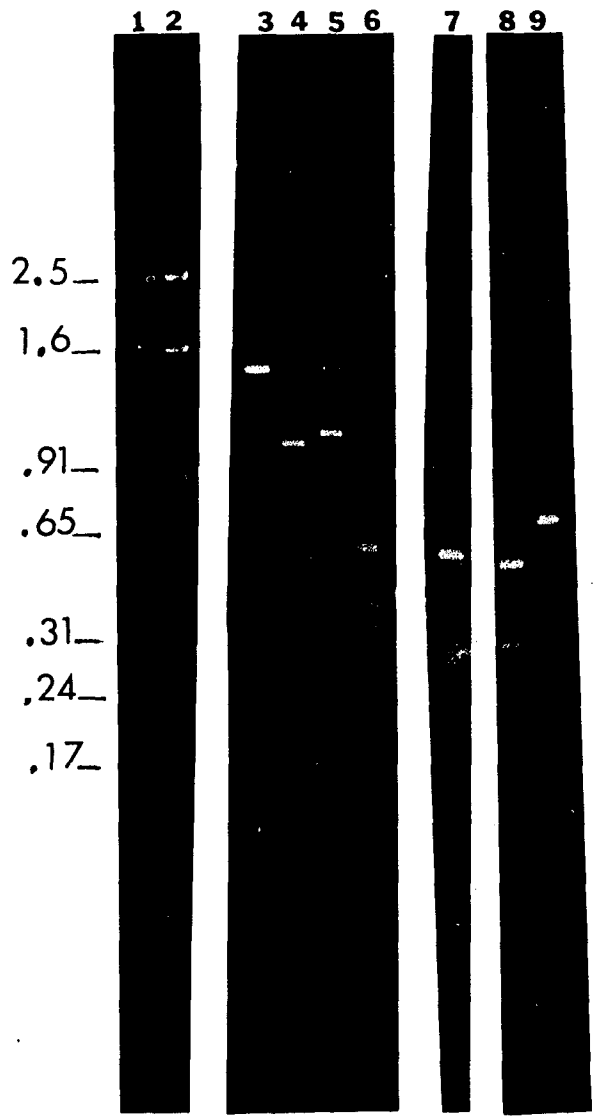

1 2
3 4 5 6
7
8 9
2.5
1.6
.91
.65
.31
.24
.17

FIGURE 8: Restriction endonuclease digests of the Bam HI insert of pPGA7 electrophoresed on a 2% agarose-TBE gel

Lanes 1 and 5 are Hae III digests of M13mp8, and were used as length standards. Lengths shown are in Kilobases. Lane 2, Hae III; Lane 3 Hpa II; Lane 4, Hae III plus Hpa II.

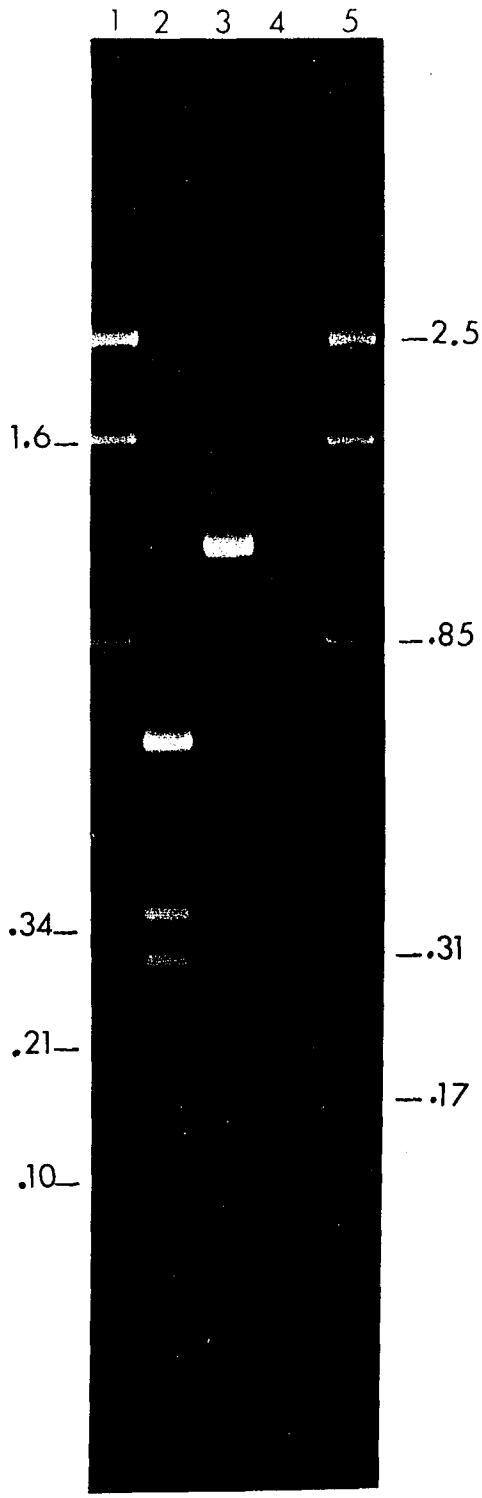

1 2 3 4 5
2.5
1.6
.85
.34
.31
.21
.17
.10

FIGURE 9: Restriction endonuclease digests of the 1.0kb and 0.42 kb Hinc II fragments of the Bam HI insert of pPGA7 electrophoresed on a 7% horizontal acrylamide-TBE gel

Lane 1, Hae III digest of the 1.0 kb fragment; Lane 2, Hpa II digest of the 1.0kb fragment; Lane 3, Hinc II digest of the Bam HI insert; Lane 4, the 0.42 kb fragment of the Hinc II digest of the Bam HI insert; Lane 5, Hae III digest of the 0.42 kb fragment; Lane 6, Hpa II digest of the 0.42 kb fragment.

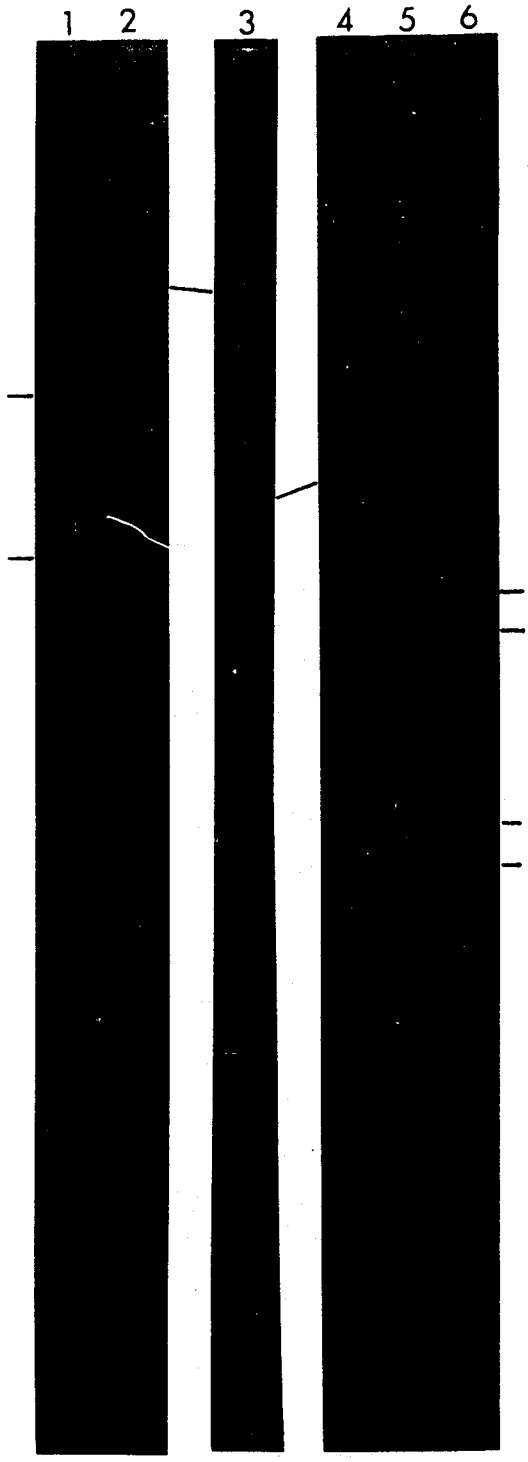

1
2
3
4
5
6

TABLE 2: Mapping of restriction endonuclease sites in pPGA7

Data from Figure 7

Fragment sizes (in kilobases)

Eco RI:	1.1	Hinc II:	1.0	Hinc II + Hae III:	0.56
	0.4		0.42		0.35
	1.5		1.42		0.31
					0.17
Eco RI + Hpa II:	0.69	Eco RI + Hae III:	0.54		1.42
	0.40		0.35		
	0.27		0.31		
(not seen)	0.06		0.17		
	1.42	(not seen)	0.05		
			1.42	Conclusions	

Conclusions

1: The 1.0 kb Bam HI-Eco RI fragment is cut by Hpa II into 0.69 kb, 0.27 kb, and 0.06 kb fragments.

2: The 0.4 kb Bam HI-Eco RI fragment is not cut by Hpa II.

3: The 0.4 kb Bam HI-Eco RI fragment is cut by Hae III into 0.35 kb, and 0.05 kb fragments. This places a Hae III site 0.35 kb from the right end of the insert.

4: One Hae III site is very near the Hinc II site.

5: This result requires that the 0.17 kb Hae III fragment (see data from figure 8) be adjacent to the Hinc II site since there is no other way for Hpa II to be located 0.01 kb from the end of this fragment.

6: The location of the 0.6 kb Hae III fragment is also confirmed by this data since this fragment is reduced to 0.54 kb by digestion with Eco RI.

TABLE 2 (continued)

Data from Figures 8 and 9

Fragment sizes (in kilobases) from Figure 8

Hae III:	Hpa II:	Hae III + Hpa II:
0.6	1.1	0.6
0.35	0.25	0.35
0.31	0.06	0.24
0.17	1.41	0.16
1.43		0.06
		1.41
		one very small fragment is missing

Conclusions from Figure 8

1: There are 3 Hae III and 2 Hpa II sites in the pPGA7 insert.

2: The 0.6 and 0.35 kb Hae III fragments are uncut by Hpa II.

3: The 0.6 kb Hpa II fragment is uncut by Hpa II.

4: The 0.31 kb Hae III fragment is cut by Hpa II into 0.24 kb and 0.06 kb fragments.

5: The 0.17 kb Hae III fragment is cut by Hpa II into 0.16 kb and 0.01 kb fragments.

6: The 0.25 kb Hpa II fragment is cut by Hae III into 0.24 and 0.005 kb fragments.

7: The 1.1 kb Hpa II fragment is cut by Hae III into 0.6 kb and 0.35 kb fragments.

Conclusions from Figure 9

1: The 1.0 kb Bam HI-Hinc II fragment is cut by Hae III but not by Hpa II.

2: Therefore, the 1.1 kb Hpa II fragment must include the 1.0 kb Bam HI-Hinc II fragment. This places one Hpa II site about 0.1 kb to the left of the Hinc II site, and places the 0.6 kb and the 0.35 kb Hae III fragments within the 1.0 kb Bam HI-Hinc II fragment.

3: The 0.42 kb Bam HI-Hinc II fragment is cut by both Hae III and Hpa II. Therefore, the 0.42 kb Bam HI-Hinc II fragment must contain the 0.31 kb and 0.17 kb Hae III fragments, and the 0.25 Kb Hpa II fragment.

4: See Figure 10 for the final location of the Hae III and Hpa II sites in the Bam HI insert of pPGA7

FIGURE 10: The restriction endonuclease map of the Bam HI insert of pPGA7

The lengths of the various restriction fragments are in kilobases. This map was determined from Figures 7, 8, and 9 and from the data in Table 2.

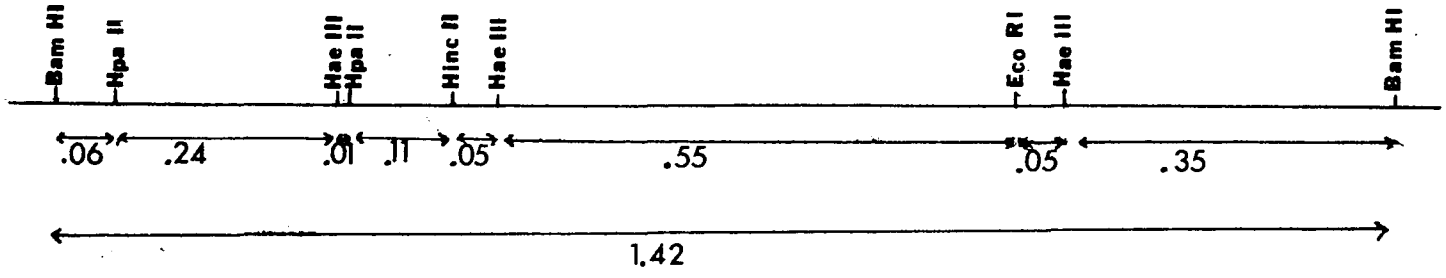

Bam HI
Hpa II
Hae III
Hpa II
Hinc II
Hae III
Eco RI
Hae III
Bam HI
.06
.24
.01
.11
.05
.55
.05
.35
1.42

FIGURE 11: Pst I digest of pPGB9 electrophoresed on a 2% agarose-TBE gel.

Lane 1 is a Pst I partial digest of pPGB9, with fragments of 5.4 kb, 2.8 kb, 2.45 kb, 0.33 kb, and 0.24 kb in length. Lane 2 is an Alu I digest of pBR322 and lane 3 is a Hind III plus Eco RI digest of lambda DNA. These were used as length standard and selected lengths are listed in kilobases.

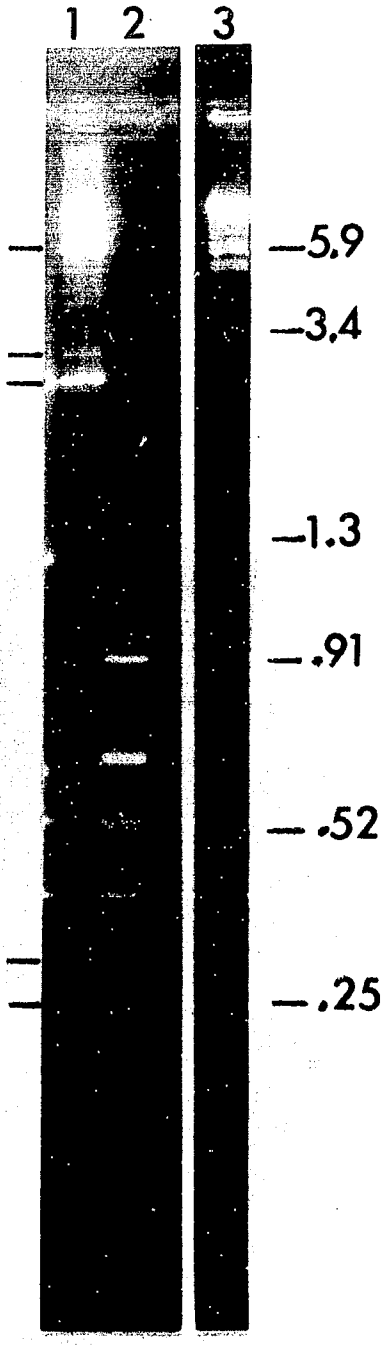
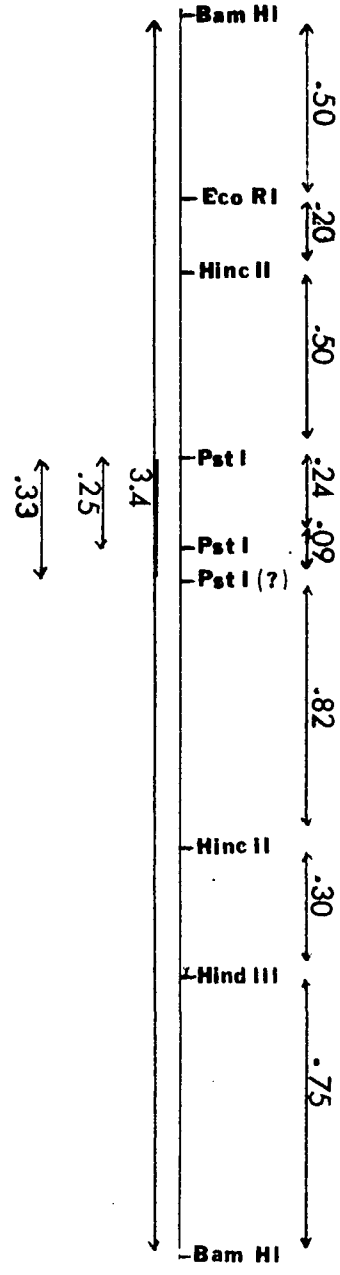

1
2
3
5.9
3.4
1.3
.91
.52
.25

FIGURE 12: Restriction endonuclease map of pPGB9

The restriction endonuclease map of pPGB9, and the location of the Pst I sites of pPGB9 as determined from Figure 11. Fragment lengths are listed in kilobases.


Bam HI
Eco RI
Hinc II
Pst I
Pst I
Pst I (?)
Hinc II
Hind III
Bam HI
.50
.20
.50
.24
.09
.82
.30
.75
3.4
.25
.33

sheared, calf thymus DNA. These hybridizations were carried out at low Cot values to select for repetitive sequences. Figures 13 and 14 show the gels and autoradiograms of three digests of the Bam HI insert of pPGA7. Figure 15 shows a gel and autoradiogram of the Pst I digest of pPGB9.

Figure 13 shows that the 0.6 kb and 0.42 kb fragments of the Hinc II plus Eco RI digest of the Bam HI insert of pPGA7 hybridize, while the 0.4 kb fragment does not. The 0.95 kb fragment is a partial digest product which includes the 0.6 kb fragment. The 1.4 kb fragment represents the entire Bam HI insert.

Figure 14 shows electrophoretic gels and autoradiograms of a Hae III digest and Hpa II digest of the Bam HI insert of pPGA7. The 1.1 kb and 0.26 kb Hpa II fragments and the 0.31 kb and 0.17 kb Hae III fragments hybridized to the radioactive probe. The 0.6 kb and 0.35 kb Hae III fragments show some hybridization but the bands on the autoradiogram are not as intense as the bands for the 0.31 kb and 0.17 kb Hae III fragments. The weaker signal is probably due to nonrepetitive sequence hybridization. These results indicate that the repeated sequence in pPGA7 is located in the 0.42 kb region shown in Figure 16a.

Figure 15 shows an electrophoretic gel and autoradiogram of the Pst I digest of pPGB9. This figure shows that the entire 3.4 kb insert, as well as the 0.24 kb Pst I fragment hybridized to the radioactive probe. The 0.09 kb fragment was too small to be detected in this experiment. This localizes the repetitious DNA to the 0.24 kb fragment (possibly including the 0.09 kb fragment) of pPGB9 as shown in Figure 16b.

FIGURE 13: Gel and autoradiogram.

A 2% agarose-TBE gel and autoradiogram of a Hinc II + Eco RI digest and a Hinc II + Hpa II digest of the Bam HI insert of pPGA7. A Southern transfer of the gel was hybridized to ^{32}P-nick translated calf thymus DNA. Lane 1 is an Alu I digest of pBR322 DNA which was used as a length standard. Lengths shown are in Kilobases. Lane 2, Hinc II + Eco RI, fragments 1.4 kb, 0.95 kb, 0.6 kb, 0.42 kb, and 0.4 kb in length. Lane 3, Hinc II + Hpa II, fragment lengths of 1.4 kb, 0.35 kb, 0.31 kb, 0.25 kb, and 0.13 kb. Lane 4 is the autoradiogram from the Southern transfer of the DNA in lanes 2 and 3.

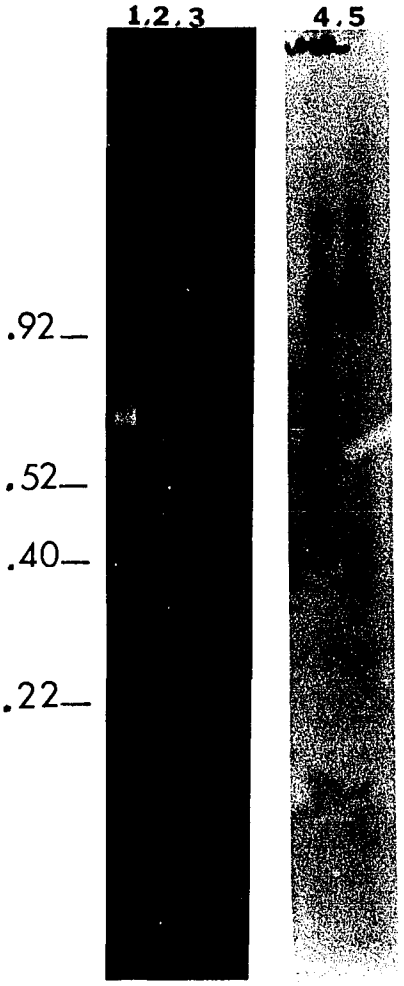

1,2,3
4,5
.92
.52
.40
.22

FIGURE 14: Gel and autoradiogram.

A 2% agarose-TBE gel and autoradiogram of restriction endonuclease digests of the Bam HI insert of pPGA7. Lane 1, Hind III + Eco RI digest of lambda DNA which was used as a length standard. Lane 4, an Alu I digest of pBR322 DNA which was used as a length standard. Lengths shown are in kilobases. Lane 2, Hae III, fragments 0.6 kb, 0.35 kb, 0.31 kb, and 0.17 kb fragments; Lane 3, Hpa II, fragments 1.4 kb, 1.1 kb, and 0.26 kb in length; Lane 5 autoradiogram of the digest in lane 2; Lane 6, autoradiogram of the digest in lane 3.

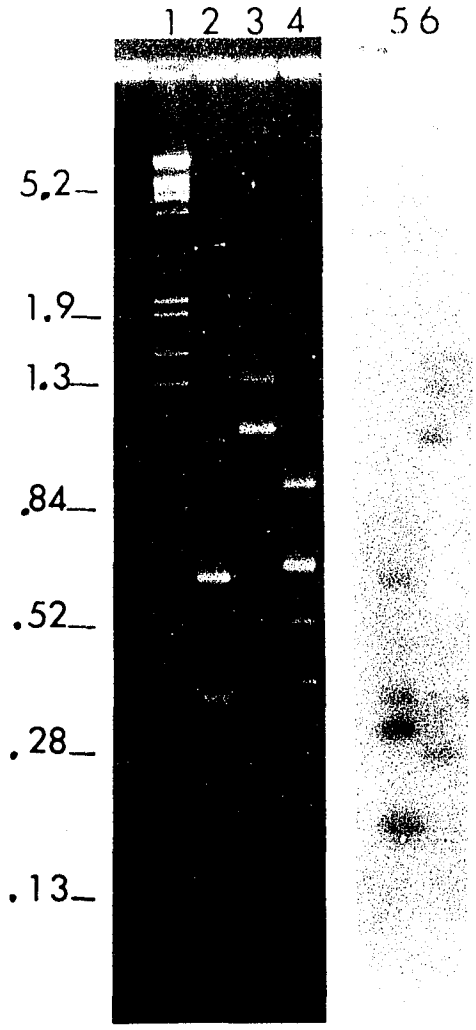

1 2 3 4
5 6
5.2
1.9
1.3
.84
.52
.28
.13

FIGURE 15: Gel and autoradiogram

A 2% agarose gel and autoradiogram of the Pst I digest of pPGB9. Lane 1, autoradiogram of the digest in lane 2. Lane 2, Pst I, fragments 5.4 kb, 2.8 kb, 2.45 kb, 0.33 kb, and 0.25 kb in length. Lane 3, Alu I digest of pBR322 used as a length standard. Indicated lengths are in Kilobases.

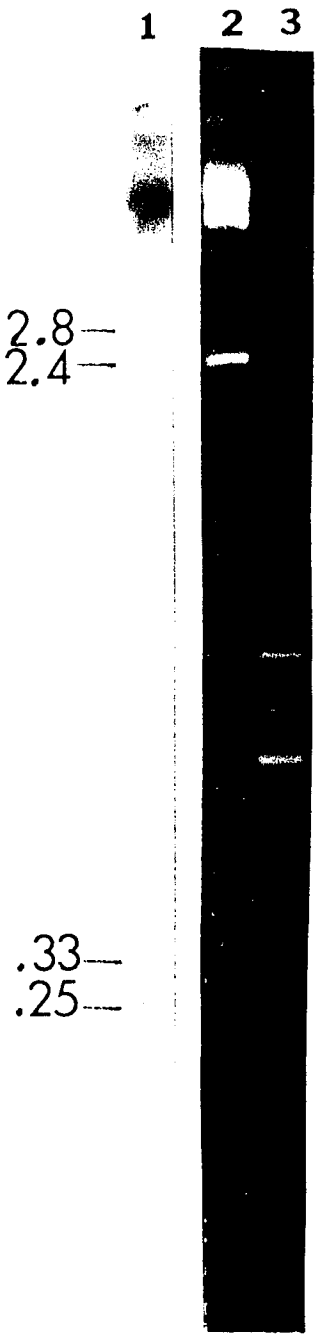
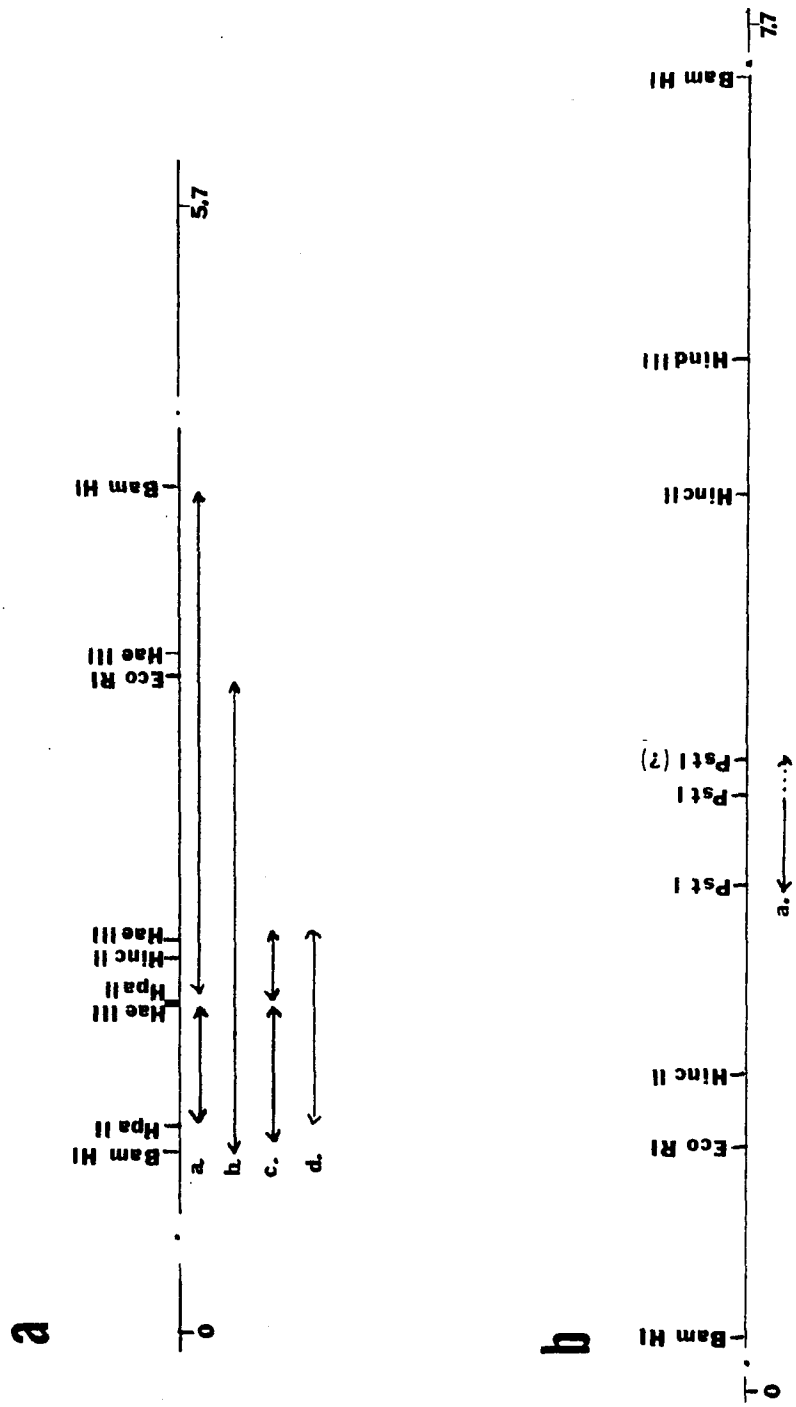

1
2
3
2.8
2.4
.33
.25

FIGURE 16: Localization of repetitious DNA in clones pPGA7 and pPGB9.

(a) Localization of the repetitious DNA in pPGA7 as determined by Figures 13 and 14. a, 0.25 Kb and 1.1 Kb; b, 1.2 Kb; c, 0.31 Kb and 0.17 Kb; d, the resultant repetitive sequence region of pPGA7 of 0.42 Kb.

(b) Localization of the repetitious DNA in pPGB9 as determined by Figure 15. a, the repetitious DNA sequence is contained in the 0.25 Kb fragment and possibly extends to the third Pst I site (dotted arrow).


a
5.7
Bam HI
Eco RI
Hae III
Hae III
Hinc II
Hpa II
Hae III
Hpa II
Bam HI
0
a.
b.
c.
d.
b
7.7
Bam HI
Hind III
Hinc II
Pst I (?)
Pst I
Pst I
Hinc II
Eco RI
Bam HI
0
a.

Cross Hybridizations between Various Repetitious DNA Containing Clones

Thirty-six recombinant plasmids, which had previously been selected to contain repetitive bovine sequences (see Materials and Methods), were electrophoresed on 1% agarose and transferred to nitrocellulose filter paper by the method of Southern (1975). This DNA was then hybridized to ^{32}P-nick translated probes. The probes used were calf thymus DNA, the Bam HI insert of pPGA7, or the Bam HI insert of pPGB9. pBR322 DNA was run on each gel as a control to show that the radioactive probe DNA did not hybridize to pBR322 DNA. Also, a plasmid containing a human Alu type repetitive sequence (BLUR-8) was hybridized to calf thymus DNA and the Bam HI insert of pPGA7. This was performed to check for homology between the human and bovine repetitive sequences.

Figures 17a and b show the gel and autoradiogram of the undigested bovine clones hybridized to ^{32}P-nick translated calf thymus DNA. Comparison of figure 17a with 17b shows that most of the clones hybridized to labelled bovine DNA at low Cot.

Figure 18 shows a gel and autoradiogram of a some of the recombinants in the bovine clone library digested with Bam HI and also hybridized to labelled calf thymus DNA. This experiment was performed by Nancy Rosen (graduate student, Iowa State University). This and other similar experiments show that 26 of the 36 recombinants containing bovine DNA inserts, hybridize to the probe. (pBR322 DNA did not hybridize under these conditions.)

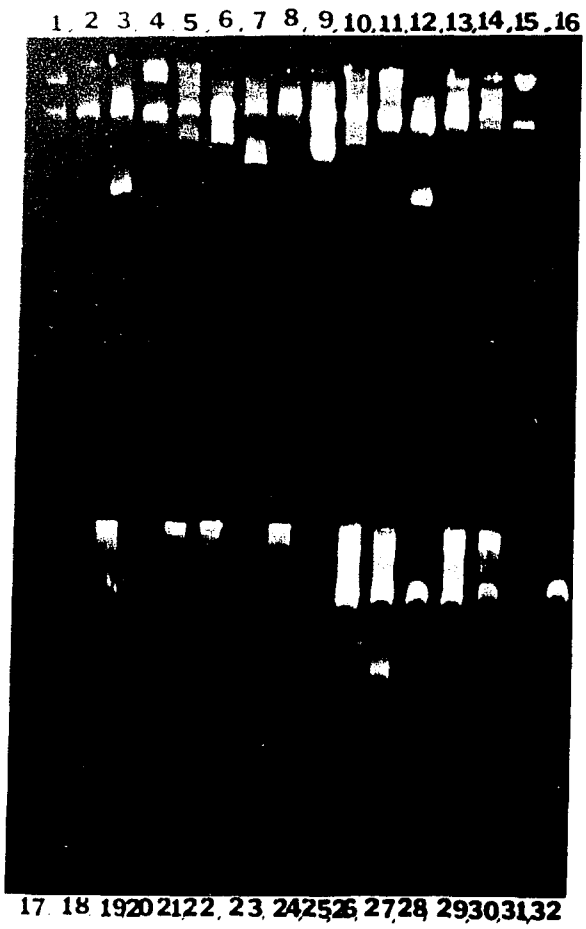
Figure 19 shows the gel and autoradiogram of Bam HI digested clones hybridized to ^{32}P-nick translated pPGA7. Since both the bovine insert and

FIGURE 17: Gel and autoradiogram of bovine recombinants

(a) Undigested DNA preps of bovine recombinants electrophoresed on 2% agarose. The following is a list of the lanes and the corresponding clones: 1, pPGN7; 2, pPGCC4; 3, pPGG1; 4, pPGB4; 5, pPGT1; 6, pPGF4; 7, pPGN9; 8, pPGR3; 9, pPGM1; 10, pPGT2; 11, pPGBB7; 12, pPGA7; 13, pPGG7; 14, pPGBB4; 15, pPGJ4; 16, Alu I digest of pBR322 (control); 17, pPGN4; 18, pPGQ7; 19, pPGH1; 20, BLUR-8; 21, pPGI7; 22, pPGS3; 23, pPGD5 24, pPGO8; 25, pPGS4; 26, pPGDD2; 27, pPGG3; 28, pPGI6; 29, pPGM9; 30, pPGB5; 31, pPGP5; 32, pPGB9.

(b) Autoradiogram of the gel in 'a'. Nick translated calf thymus DNA was the hybridization probe. The primed lane numbers of the autoradiogram correspond to the unprimed lane numbers in 'a'. The clones which hybridize are: pPGCC4, pPGB4, pPGT1, pPGF4, pPGN9, pPGM1, pPGT2, pPGBB7, pPGA7, pPGG7, pPGBB4, pPGJ4, pPGN4, pPGH1, pPGI7, pPGD5, pPGO8. pPGG3, pPGM9, pPGB5, pPGB9.

a


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32

b

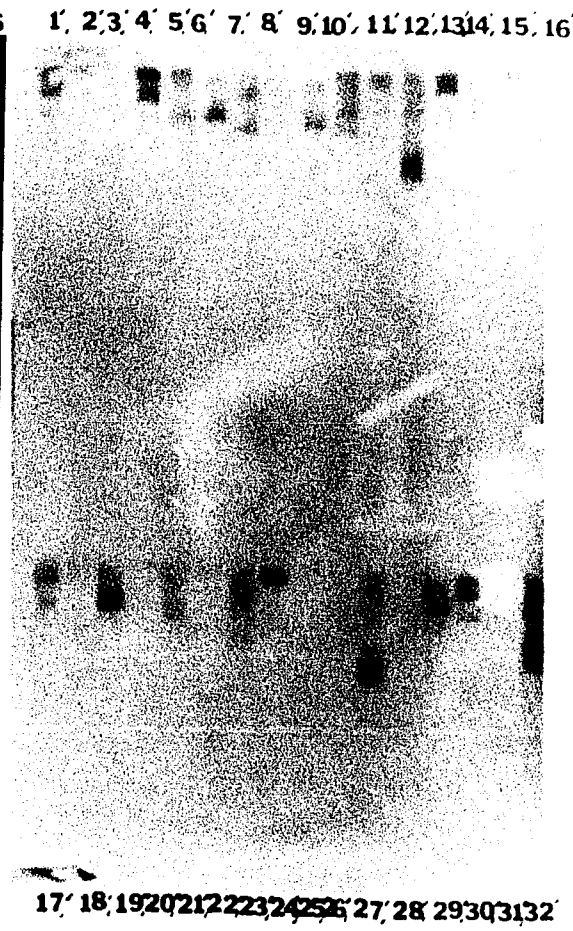

1' 2' 3' 4' 5' 6' 7' 8' 9' 10' 11' 12' 13' 14' 15' 16'
17' 18' 19' 20' 21' 22' 23' 24' 25' 26' 27' 28' 29' 30' 31' 32'

FIGURE 18: Gel and autoradiogram of bovine recombinants.

(a) Bam HI digested bovine recombinants electrophoresed on 2% agarose. The following is a list of the lane numbers and the corresponding clones: 1, Hind III plus Eco RI digest of lambda DNA (control); 2, pPGD5; 3, pPGS3; 4, pPGB4; 5, pPGCC9; 6, pPGN4; 7, pPGM3; 8, pPGT2; 9, pPGA7; 10, pPGBB7; 11, pPGS4; 12, Alu I digest of pBR322 DNA (control).

(b) Autoradiogram of the gel in 'a'. Nick translated calf thymus DNA was the hybridization probe. The primed lane numbers in the autoradiogram correspond to the unprimed lane numbers in 'a'. The clones which hybridize are: pPGBB7, pPGA7, pPGT2, pPGN4, pPGCC9, pPGB4, pPGD5.

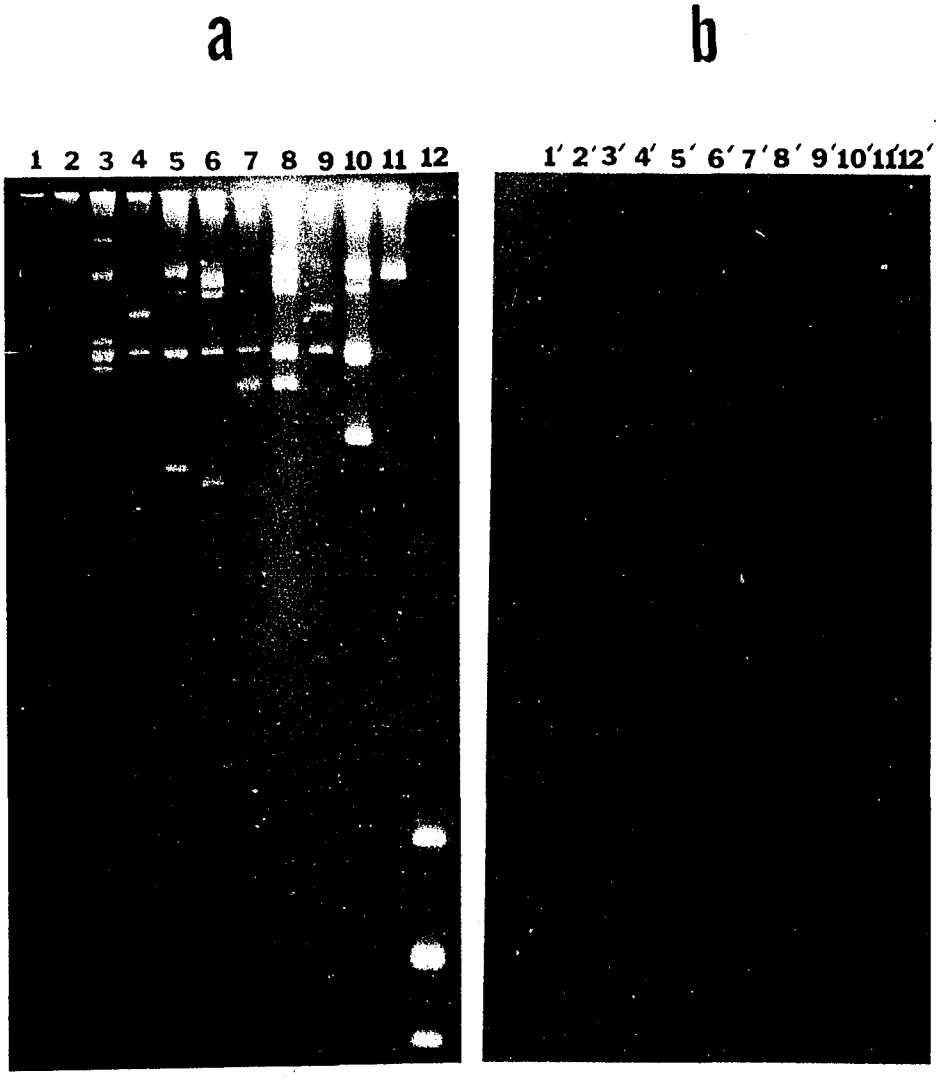
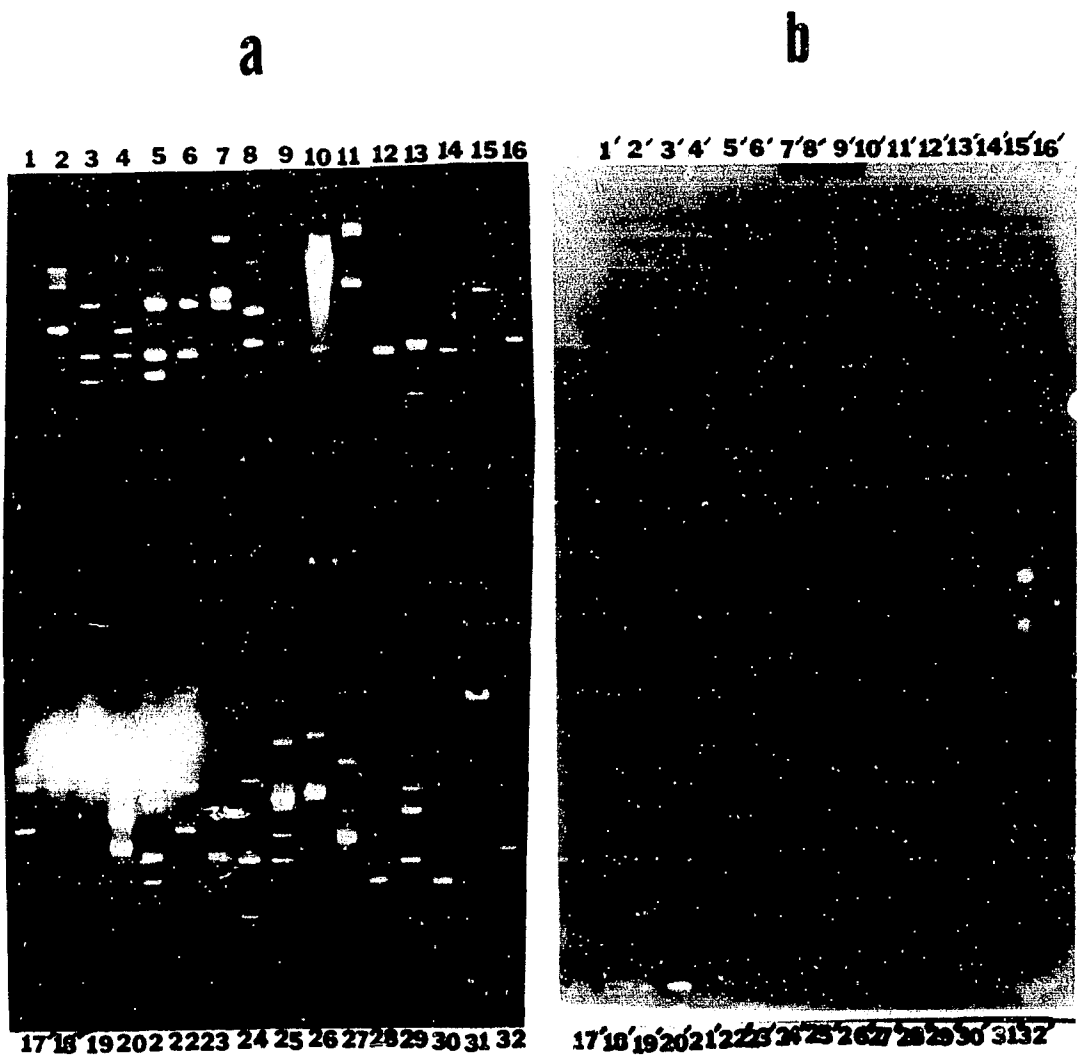

a
b
1 2 3 4 5 6 7 8 9 10 11 12
1′ 2′ 3′ 4′ 5′ 6′ 7′ 8′ 9′ 10′ 11′ 12′

FIGURE 19: Gel and autoradiogram of bovine recombinants.

(a) Bam HI digested bovine recombinants electrophoresed on 1% agarose. The following is a list of the lanes and corresponding clones: 1, nothing; 2, pPGA7; 3, pPGB9; 4, pPGB5; 5, pPGF4; 6, pPGG6; 7, pPGBB4; 8, pPGCC9; 9, pPGP5; 10, pPGBB7; 11, pPGN7; 12, pPGD5; 13, pPGQ7; 14, pPGQ9; 15, pPGB4; 16, pPGB3; 17, pPGA7; 18, pPGN9; 19, pPGS4; 20, pPGS3; 21, pPGT1; 22, pPGI6; 23, pPGG1; 24, pPGM3; 25, pPGH1; 26, pPGM9; 27, pPGG9; 28, pPGR3; 29, pPGG7; 30, pPGG3; 31, pPGF2; 32, pPGB3.

(b) Autoradiogram of the gel in 'a'. Nick translated clone pPGA7 was used as the hybridization probe. The primed lane numbers of the autoradiogram correspond to the unprimed lane numbers of 'a'. The clones which cross hybridize are: pPGA7, pPGP5, pPGQ7, pPGS3, pPGI6, pPGG1, and pPGM3.


a
b
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
1' 2' 3' 4' 5' 6' 7' 8' 9' 10' 11' 12' 13' 14' 15' 16'
17' 18' 19' 20' 21' 22' 23' 24' 25' 26' 27' 28' 29' 30' 31' 32'

the pBR322 DNA are labelled, every clone hybridized. In this experiment, some of the clones were not digested by Bam HI and therefore could not be scored. Plasmids which could be scored as positive for repetitious sequences are: pPGB9, pPGA7, pPGBB7, pPGD9, pPGS4, pPGF6, pPGG1, and pPGM3. To clarify this result, the Bam HI insert of pPGA7 was electroeluted, labelled by nick translation, and hybridized to the undigested clones (Figure 20). From this figure, it is clear that most of the clones hybridize to the labelled probe. Again, the human clone, BLUR-8, and the pBR322 DNA did not hybridize.

A similar experiment was performed using the labelled Bam HI insert of pPGB9. However, due to its similar size, the insert (uncontaminated by pBR322 DNA) of pPGB9 was impossible to purify. To overcome this problem, Southern blots were prehybridized with unlabelled, Bam HI digested pBR322 DNA. After 24 hours, radioactive pPGB9 Bam HI insert was added and hybridized for an additional 8-16 hours. Figures 21a and b show the results of this hybridization. Most of the clones show weak hybridization and there was no hybridization to the pBR322 DNA.

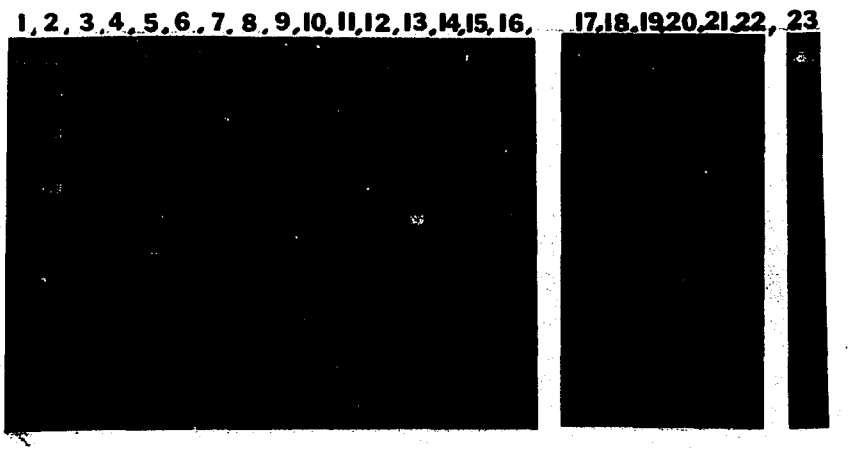
Table 3 summarizes the cross hybridization data derived from figures 17-21, as well as from several other gels not shown. It is clear that clones pPGA7 and pPGB9 cross hybridize. Also, 22 of the 26 clones which hybridize to calf thymus DNA, hybridize to the Bam HI insert of pPGA7 or pPGB9. This indicates that most of the short interspersed repeated DNA sequences in the bovine genome have closely related nucleotide sequences. It is also evident from this table that pBR322 DNA, lambda DNA, and BLUR-8 DNA do not hybridize to the radioactive probe DNAs.

FIGURE 20: Gel and autoradiogram of bovine recombinants.

(a) Bovine recombinant clones electrophoresed on 0.8% agarose. The following is a list of the lanes and corresponding clones: 1, pPGA7; 2, pPGB9; 3, pPGF4; 4, pPGG6; 5, pPGB5; 6, pPGQ7; 7, pPGCC9; 8, pPGBB4; 9, pPGN4; 10, pPGN9; 11, pPGQ9; 12, pBR322 (control); 13, pPGG3; 14, pPGT1; 15, pPGM1; 16, pPGG1; 17, pPGS3; 18, pPGH1; 19, pPGG7; 20, pPGM9; 21, pPGR3; 22, pPGB3; 23, BLUR-8.

(b) Autoradiogram of the gel in 'a'. Nick translated Bam HI insert of pPGA7 was used as the hybridization probe. The primed lane numbers of this autoradiogram correspond to the unprimed lane numbers in 'a'. The clones which hybridize are: pPGA7, pPGB9, pPGF4, pPGG6, pPGB5, pPGQ7, pPGCC9, pPGN4, pPGN9, pPGQ9, pPGG3, pPGT1, pPGM1, pPGS3, pPGH1, pPGG7, pPGM9, pPGR3, and pPGB3.

a


1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,
17, 18, 19, 20, 21, 22, 23

b

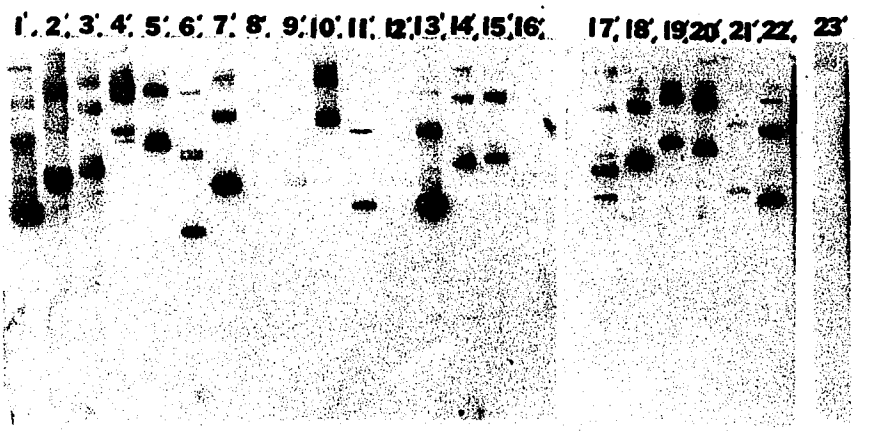
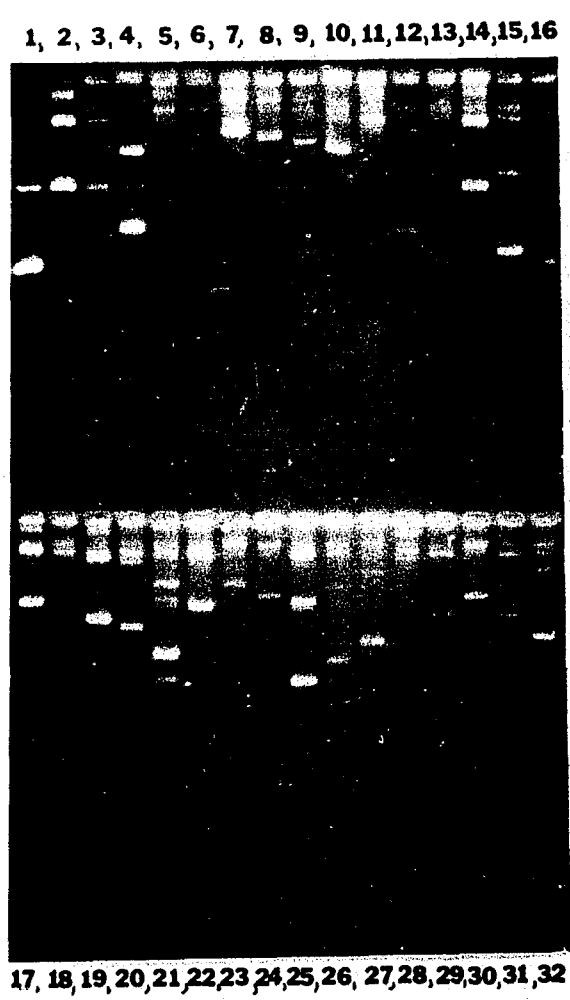

1', 2', 3', 4', 5', 6', 7', 8', 9', 10', 11', 12', 13', 14', 15', 16',
17', 18', 19', 20', 21', 22', 23'

FIGURE 21: Gel and autoradiogeam of bovine recombinants

(a) Bovine recombinant clones electrophoresed on 2% agarose. The following is a list of the lanes and corresponding clones: 1, pBR322 (control); 2, pPGB9; 3, pPGF4; 4, pPGA7; 5, pPGJ4; 6, pPGF2; 7, pPGBB7; 8, pPGN9; 9, pPGG7; 10, pPGM9; 11, pPGCC9; 12, pPGG1; 13, pPGM3; 14, pPGT2; 15, pPGQ7; 16, pPGP5; 17, pPGBB4; 18, pPGS4; 19, pPGT1; 20, pPGR3; 21, pPGI6; 22, pPGDD2; 23, pPGN7; 24, pPGB4; 25, pPGS3; 26, pPGG3; 27, pPGD5; 28, pPGO8; 29, pPGI7; 30, pPGH1; 31, pPGM1; 32, pPGN4.

(b) The autoradiogram of the gel in 'a'. Nick translated Bam HI insert of pPGB9 was used as the hybridization probe. Unlabelled Bam HI cut pBR322 had been hybridized to the filter prior to addition of the radioactive probe. The primed numbers of the autoradiogram correspond to the unprimed numbers in 'a'. The clones which hybridized are: pPGA7, pPGB4, pPGB9, pPGD5, pPGF2, pPGG7, pPGH1, pPGI7, pPGM9, pPGN7, pPGN9, pPGO8, pPGT1, pPGT2, and pPGBB7.

a


1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16
17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32

b

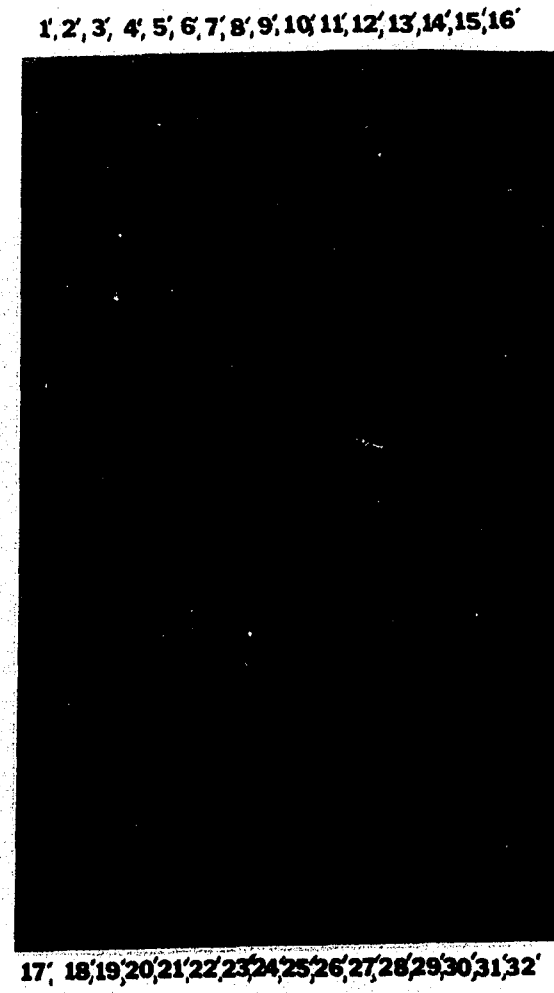

1', 2', 3', 4', 5', 6', 7', 8', 9', 10', 11', 12', 13', 14', 15', 16'
17', 18', 19', 20', 21', 22', 23', 24', 25', 26', 27', 28', 29', 30', 31', 32'

TABLE 3: Summary of cross hybridization experiments

Clone	Hybridize to labelled calf thymus DNA	Hybridize to labelled Bam HI insert of pPGA7	Hybridize to labelled Bam HI insert of pPGB9
pPGA7	+	+	+
pPGB3	+	+	ND
pPGB4	+	+	+
pPGB5	+	+	ND
pPGB9	+	+	+
pPGD5	+	ND	+
pPGF2	+	ND	+
pPGF4	+	+	-
pPGG1	+	-	-
pPGG3	+	+	-
pPGG6	+	+	ND
pPGG7	+	+	+
pPGH1	+	+	+
pPGI7	+	+	+
pPGJ4	+	ND	-
pPGM1	+	+	-
pPGM9	+	+	+
pPGN4	+	-	-
pPGN7	+	+	+
pPGN9	+	+	+
pPGO8	+	+	+
pPGT1	+	+	+
pPGT2	+	ND	+
pPGBB4	+	-	-
pPGBB7	+	-	+
pPGCC9	+	+	-
pBR322	-	-	-
BLUR-8	-	-	ND
Lambda DNA	-	ND	ND

Cross hybridization of ^{32}P-labelled calf thymus DNA, the Bam HI insert of pPGA7, and the Bam HI insert of pPGB9 to 26 recombinant clones containing bovine repetitive sequences. pBR322, and lambda DNA are controls. BLUR-8 is a clone of a member of the human Alu sequence family of short interspersed repeated sequences. In the above table, 22 of the 26 clones cross hybridize to the Bam HI inserts of either pPGA7 or pPGB9.

S1 Nuclease Determined Sizes of Bovine Repetitive Sequences

Repetitive sequence lengths in total bovine DNA

Most interspersed repetitive sequences in bovine DNA have been shown by electron microscopy to be about 300 nucleotides in length (Mayfield, et al., 1980). S1 nuclease digestion of hybrids is an independent method of determining repetitive sequence length.

Bovine DNA from calf thymus was sheared using a French press to an average length of 500-600 nucleotides (Figure 22). A portion of this DNA was heat denatured, and then allowed to self-hybridize to low Cot values so that only repetitive DNA could hybridize. These samples were then digested with S1 nuclease. The S1-resistant, double stranded DNA was fractionated on 2% agarose (Figure 23). The lengths of the S1 resistant DNAs were determined by comparison with known lengths of DNA. Figure 23, lane 2 shows the 3 distinct size classes of double stranded DNA. These repetitive sequence lengths are: 300 bp, 400 bp, and 600 bp.

Determination of the repetitious sequence length in pPGA7 and pPGB9

The length of the repetitive DNA in clones pPGA7 and pPGB9 was determined by S1 nuclease digestion of clone DNA-calf thymus DNA hybrids. Resolution of the resultant S1-resistant fragments was by gel electrophoresis.

The Bam HI inserts of pPGA7 and pPGB9 were electroeluted from 7% acrylamide gels and nick translated. In each hybridization experiment, 0.01-0.1 ugs of DNA (SA: 10^7 - 10^8 cpm/ug DNA) were denatured and mixed with a 100-200 fold excess of sheared, denatured calf thymus DNA. The hybridizations were carried out in 0.15 M NaCl at 60^{o}C for 2 hours.

FIGURE 22: Determination of the average length of calf thymus after shearing with a French press

The DNA was put through the French press three times at 10,000-20,000 psi. Lane 1 is a Hind III plus Eco RI digest of lambda DNA used as a length standard. Lane 6 is an Alu I digest of pBR322 DNA used as a length standard. Selected lengths are in kilobases. Lanes 2 and 3 are the calf thymus DNA before shearing. Lanes 4 and 5 are the same DNA after shearing. The average size of the DNA in lanes 4 and 5 is 500 nucleotides.

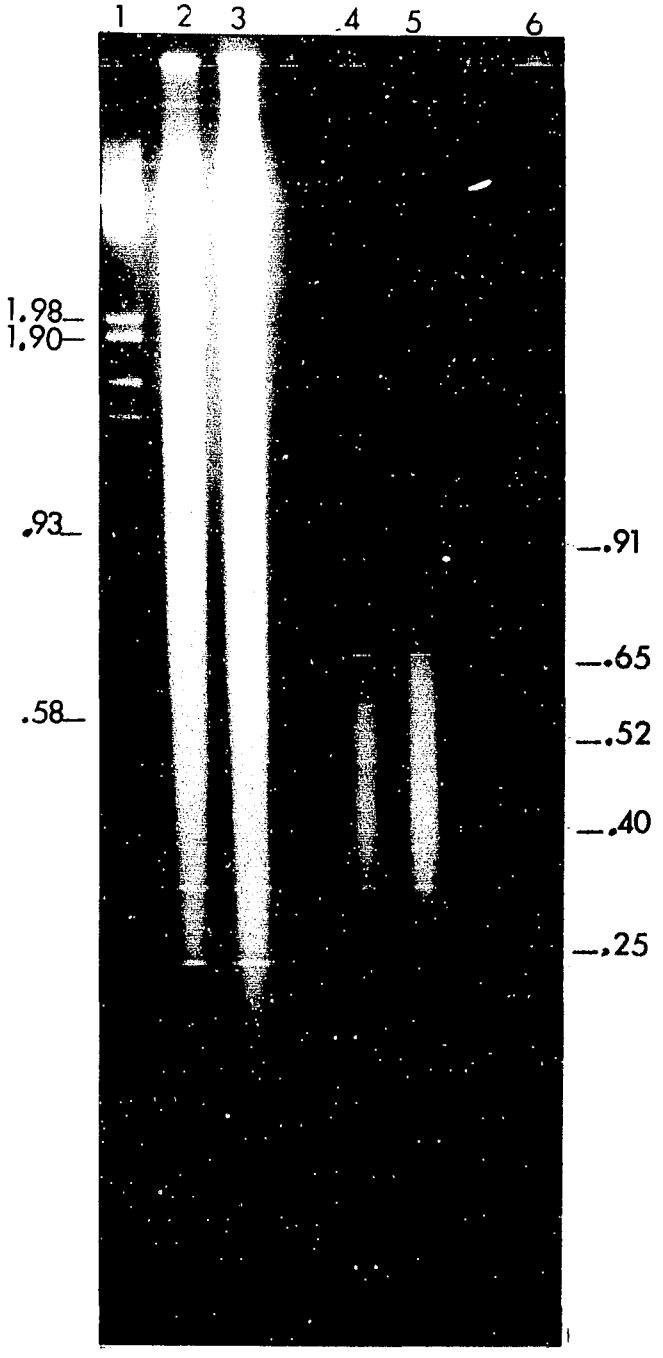
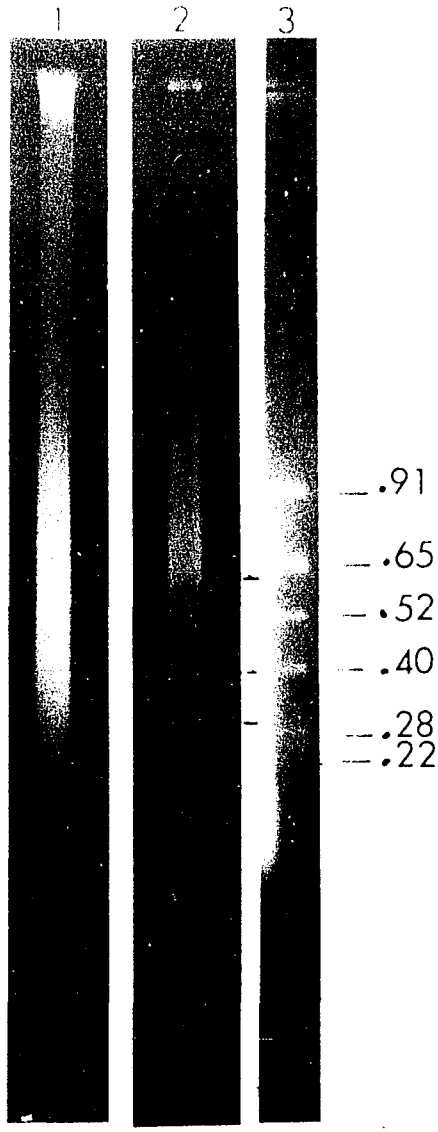

1
2
3
4
5
6
1.98
1.90
.93
.58
.91
.65
.52
.40
.25

FIGURE 23: S1 resistant, repetitive bovine DNA hybrids

Total bovine DNA was allowed to self-hybridize to low Cot, digested with S1 nuclease, and then fractionated on 2% agarose. Lane 1 is the self-hybridized, sheared calf thymus DNA before treatment with S1 nuclease. Lane 2 is the same DNA after treatment with S1 nuclease. Lane 3 is an Alu I digest of pBR322 which was used as a length standard. Lengths shown are in kilobases. The three length classes of S1 resistant DNAs are: 300 bp, 400 bp, and 600 bp.


1
2
3
.91
.65
.52
.40
.28
.22

Controls containing only radioactive tracer DNA were also performed.

After ethanol precipitation, samples were incubated for 1 hour at 37°C either with or without S1 nuclease. All samples were then electrophoresed on 2% agarose, the gel covered with Saran Wrap, and autoradiographed at room temperature. Figures 24a and b show the gel and autoradiogram for this experiment using clone pPGA7. Figures 25a and b show the gel and autoradiogram for this experiment using clone pPGB9.

By comparing Figure 24a lanes 2, 4, 6, and 8 with lanes 10-15, it is evident that much of the DNA was digested by S1 nuclease. Figure 24b, lanes 2′, 4′, 6′, and 8′ show that the radioactive tracer DNA is present in all length classes of DNA. Lanes 10′-15′ show that there is a length class of DNA which is resistant to digestion by S1 nuclease. These hybrid DNAs are 120-130 base pairs in length and are present in both control (tracer only) and experimental (tracer plus driver) lanes.

Figures 25a and b show the results of a similar experiment using labelled Bam HI insert of pPGB9. The result was very similar to that using pPGA7 shown in Figures 24a and b.

These results suggest that there is an internal region of homology (120-130 bp long) in the repetitive DNA sequences in both pPGA7 and pPGB9.

Percent of the Bovine Genome Complimentary to pPGA7 and pPGB9

Hybridizations were performed between calf thymus DNA and subclones of the Bam HI inserts of pPGA7 and pPGB9. These inserts were subcloned into M13mp8 (see DNA Sequencing) and are named M13mp8BA7 and M13mp8BB9. The hybrids were then treated with S1 nuclease and fractionated by

FIGURE 24: Determination of the size of the repetitious DNA in pPGA7.

(a) Nick translated Bam HI insert of pPGA7 hybridized with or without excess calf thymus DNA. Samples were treated with or without S1 nuclease, and before or after hybridization. Samples were fractionated of 2% agarose. The following is a list of lanes and the corresponding samples: 1, Hind III plus Eco RI digest of lambda DNA (control); 2, pPGA7 DNA, before hybridization, no S1 nuclease; 3, nothing; 4, pPGA7 plus calf thymus DNA, before hybridization, no S1 nuclease; 5, Alu I digest of pBR322 (control); 6, pPGA7 DNA, hybridized, no S1 nuclease; 7, nothing; 8, pPGA7 plus calf thymus DNA, hybridized, no S1 nuclease; 9, Alu I digeat of pBR322 (control); 10, pPGA7 DNA, before hybridization, 70 units S1 nuclease; 11, pPGA7 DNA, hybridized, 70 units S1 nuclease; 12, pPGA7 DNA, hybridized, 70 units S1 nuclease; 13, pPGA7 plus calf thymus DNA, before hybridization, 70 units S1 nuclease; 14, pPGA7 plus calf thymus DNA, hybridized, 70 units S1 nuclease; 15, pPGA7 plus calf thymus DNA, hybridized, 70 units S1 nuclease; 16, Alu I digest of pBR322 (control). Lengths shown are in kilobases.

(b) Autoradiogram of the gel in 'a'. The primed numbers of the autoradiograph correspond to the unprimed numbers in 'a'. The size of the S1 resistant hybrids in lanes 10'-15' is 120-130 nucleotides in length.

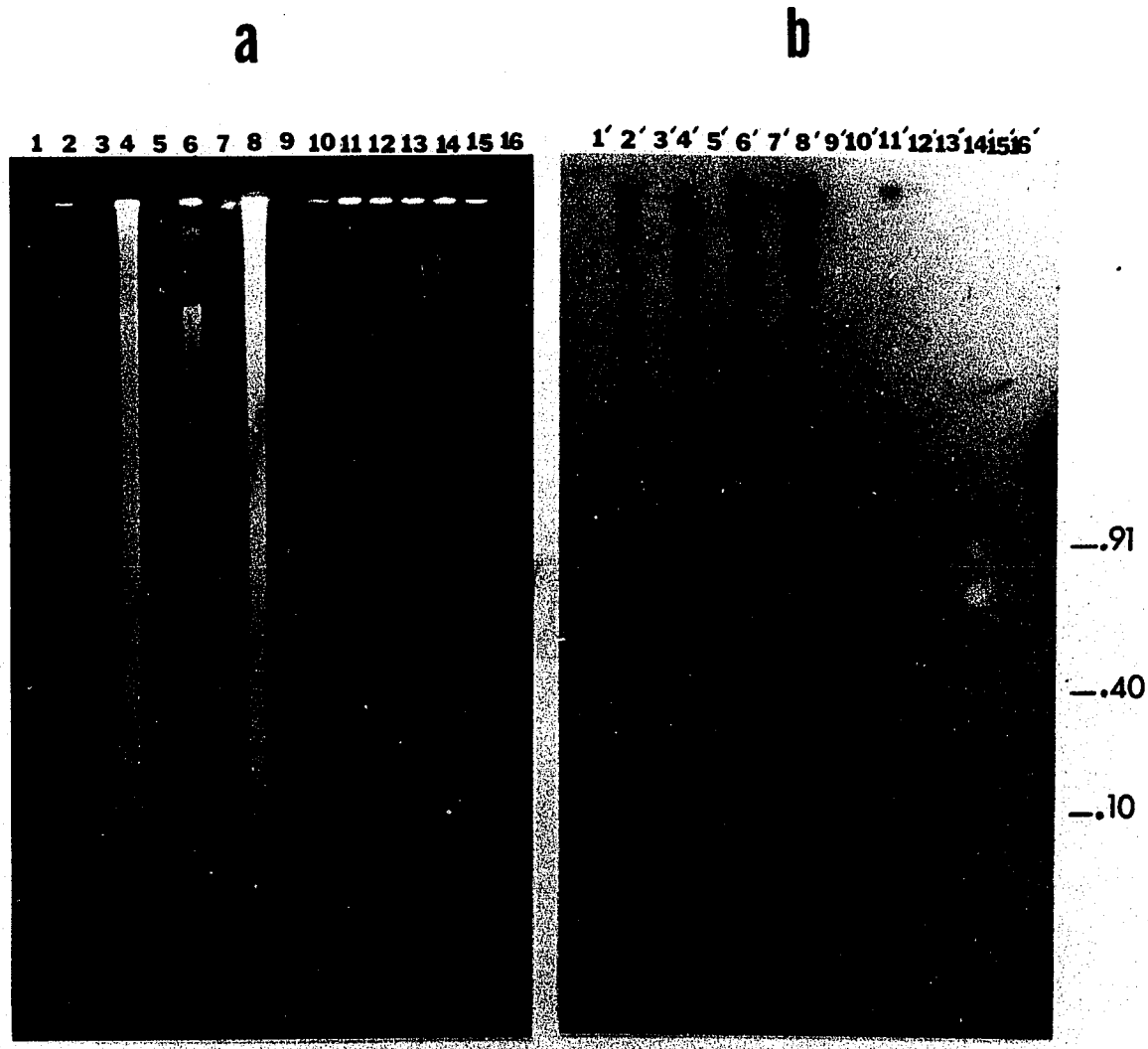
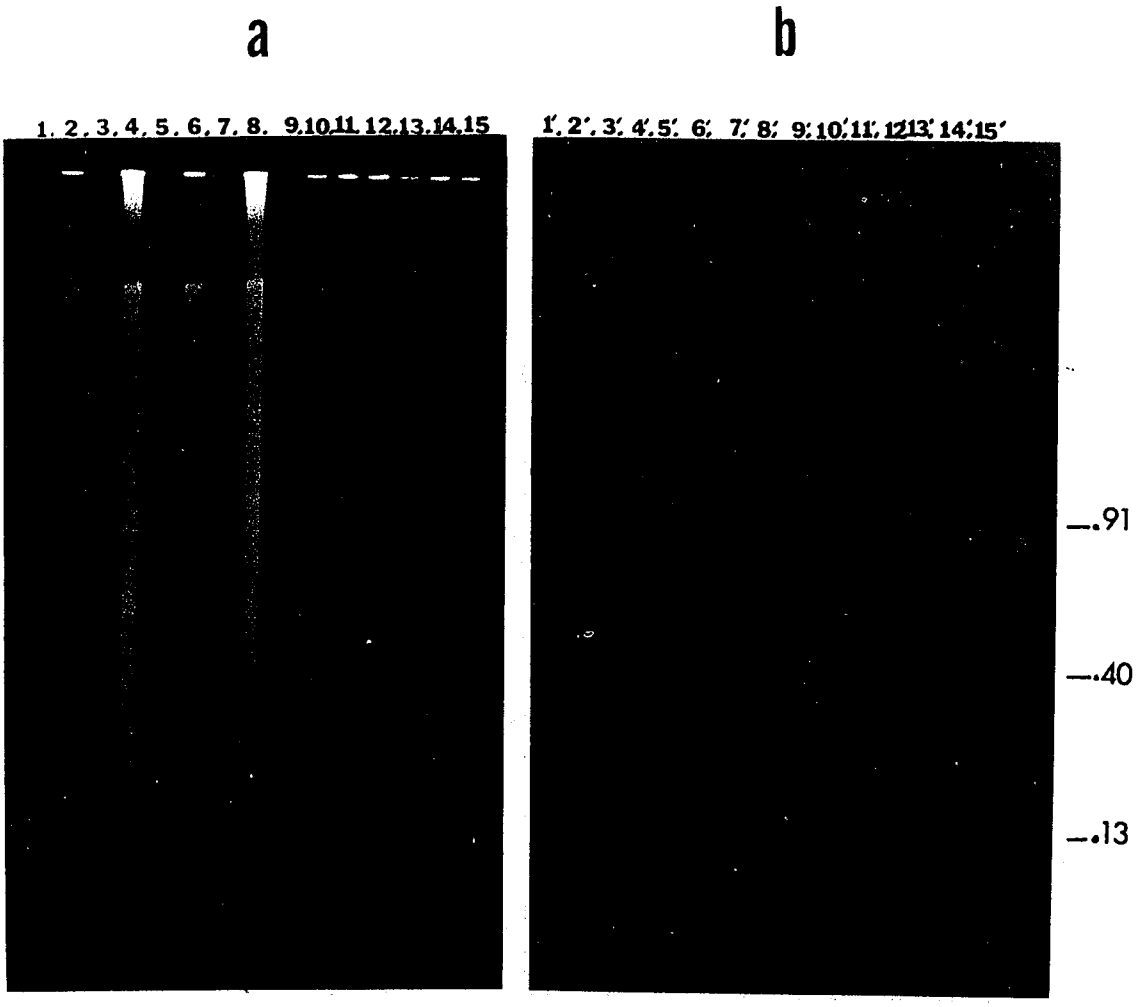

a
b
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
1' 2' 3' 4' 5' 6' 7' 8' 9' 10' 11' 12' 13' 14' 15' 16'
.91
.40
.10

FIGURE 25: Determination of the size of the repetitious DNA in pPGB9.

(a) Nick translated Bam HI insert of pPGB9 hybridized with or without excess calf thymus DNA. Samples were treated with or without S1 nuclease, and before or after hybridization. Samples were fractionated of 2% agarose. The following is a list of lanes and the corresponding samples: 1, Hind III plus Eco RI digest of lambda DNA (control); 2, pPGB9 DNA, before hybridization, no S1 nuclease; 3, nothing; 4, pPGB9 plus calf thymus DNA, before hybridization, no S1 nuclease; 5, Alu I digest of pBR322 (control); 6, pPGB9 DNA, hybridized, no S1 nuclease; 7, nothing; 8, pPGB9 plus calf thymus DNA, hybridized, no S1 nuclease; 9, Alu I digeat of pBR322 (control); 10, pPGB9 DNA, before hybridization, 70 units S1 nuclease; 11, pPGB9 DNA hybridized, 70 units S1 nuclease; 12, pPGB9 DNA, hybridized, 70 units S1 nuclease; 13, pPGB9 plus calf thymus DNA, before hybridization, 70 units S1 nuclease; 14, pPGB9 plus calf thymus DNA, hybridized, 70 units S1 nuclease; 15, pPGB9 plus calf thymus DNA, hybridized, 70 units S1 nuclease; 16, Alu I digest of pBR322 (control).

(b) Autoradiogram of the gel in 'a'. The primed numbers of the autoradiograph correspond to the unprimed numbers in 'a'. The size of the S1 resistant hybrids in lanes 10'-15' is 120-130 nucleotides in length.


a
b
1. 2. 3. 4. 5. 6. 7. 8. 9.10.11 12.13.14.15
1'. 2'. 3'. 4'. 5'. 6'. 7'. 8'. 9'. 10'. 11'. 12'13'. 14'. 15'
—.91
—.40
—.13

hydroxylapatite chromatography.

Trace amounts of ^{32}P-nick translated calf thymus DNA were hybridized with a 10-200 fold excess of unlabelled M13mp8BA7, M13mp8BB9, or M13mp8BA7 plus M13mp8BB9. Cot values of 10 and 100 were chosen to insure that all repetitive DNA would hybridize. Also, these Cot values are at least one order of magnitude below the $Cot_{1/2}$ for hybridization of bovine single copy sequences. A control sample of undriven, labelled calf thymus DNA was hybridized in parallel and used to correct for any self-hybridization of the tracer DNA. A second control was used to correct for the amount of DNA in hybrid form after heat denaturation but before hybridization.

All the samples were digested with S1 nuclease, fractionated on hydroxylapatite columns, and the amount of double stranded DNA determined.

The fractions of the bovine DNA which hybridized to M13mp8BA7 or to M13mp8BB9 were calculated as follows:

Let Y = the fraction of the labelled DNA which is double stranded following hybridization with excess clone DNA

Let Z = the fraction of the labelled DNA which is double stranded following hybridization in the absence of clone DNA

Let W = the fraction of the labelled DNA which is double stranded at zero time of hybridization with excess clone DNA

Let X = the fraction of the labelled DNA which is double stranded at zero time of hybridization without clone DNA

The fraction of the DNA which hybridized during the undriven reaction (without added clone DNA) is given by:

$$A = Z - X.$$

The calculated fraction of the labelled DNA which has hybridized to the clone DNA is given by:

$$R = W - A.$$

Since the M13 cloned DNA is single stranded and contains only one of the two sense strands of double stranded DNA, the fraction of the bovine DNA which contains repetitive sequences is twice R. Table 4 lists these values for 10 experiments. The average value of R for the Cot 10 reactions is 3.26% ± 1.71% and the average value of R for the Cot 100 reactions is 2.88% ± 0.68%. These numbers are not significantly different. This shows that the reaction was completed by a Cot of 10, and allows the data from the Cot 10 and Cot 100 reactions to be combined. The average value of R for M13mp8BA7 is 2.73%; the average value for M13mp8BB9 is 4.53%; and the average value for the combined M13mp8BA7 and M13mp8BB9 experiment is 2.66%. Since the high value of R for M13mp8BB9 is due to the single Cot 10 value, and since the M13mp8BA7 plus M13mp8BB9 values are no higher than that for M13mp8BA7 alone, it was concluded that M13mp8BA7 and M13mp8BB9 were hybridizing to the same bovine sequences. Therefore, it is legitimate to calculate an overall average for all ten experiments. This value of R is 3.07% ± 1.32% which indicates that 6.14% ± 2.64% of the bovine genome consists of repetitive sequences which will cross hybridize with M13mp8BA7 and M13mp8BB9.

DNA Sequencing

DNA sequencing was performed by both the Maxam-Gilbert and dideoxy sequencing methods. This section will first discuss the subcloning and

TABLE 4: Percent of the bovine genome represented by the repeated sequence in clones pPGA7 and pPGB9

Sample	Cot	% Hybrid(Y)	% Hybrid - 0 time(W)	% Hybrid - Undriven(A) =	Corrected % Hybrid(R)	2R
M13mp8BA7 + CT DNA	10	4.50	1.19	1.80	1.51	3.02
M13mp8BA7 + CT DNA	10	3.74	1.74	0.43	1.57	3.14
M13mp8BA7 + CT DNA	10	6.80	1.90	0.90	4.00	8.00
M13mp8BB9 + CT DNA	10	9.10	1.40	1.80	5.90	11.80
M13mp8BA7 + M13mp8BB9 + CT DNA	10	6.70	1.60	1.80	3.30	6.60
M13mp8BA7 + CT DNA	100	7.25	1.19	2.29	3.77	7.54
M13mp8BA7 + CT DNA	100	5.34	1.17	1.20	2.40	4.80
M13mp8BA7 + CT DNA	100	13.90	1.90	8.90	3.10	6.20
M13mp8BB9 + CT DNA	100	6.85	1.40	2.29	3.16	6.32
M13mp8BA7 + M13mp8BB9 + CT DNA	100	5.90	1.60	2.29	2.01	4.02

This is a list of the hybridization experiments with trace amounts of nick translated calf thymus DNA hybridized to excess M13mp8BA7, M13mp8BB9, or M13mp8BA7 plus M13mp8BB9 viral DNAs. From the calculations described in the Results section, the repetitive sequence cloned in pPGA7 and pPGB9 composes 6.14% ± 2.64%.

sequencing of pPGA7 and then pPGB9.

Subcloning and sequencing of the Bam HI insert of pPGA7

The bovine clone pPGA7, the RF form of M13mp8, and the RF form of M13mp9 were digested with Bam HI. The digested DNAs were combined and ligated at 14^{o}C for 8 hours in the presence of 0.2 units of T4 DNA ligase. E. coli strain JM103 was then transformed with the ligation mixture and plated on YT indicator plates. Six white plaques were picked and the viral DNAs screened using the DIGE test. This gel is shown in Figure 26. Lanes 1, 5, 6, and 10 are controls. Lanes 2, 3, 4, 7, 8, and 9 contain the M13 recombinants: M13mp83, M13mp84, M13mp85, M13mp97, M13mp98, and M13mp99. Each of these recombinant M13 DNAs contains a DNA insert similar in size to the Bam HI insert of pPGA7.

Other M13mp8 and M13mp9 clones containing the Bam HI insert of pPGA7 were screened in the same manner. These were hybridized to M13mp85 and M13mp98 using the C test. This result is shown in Figure 27. Recombinant clones in lanes 2, 5, 8, and 9 hybridize to M13mp85, and those in lanes 14, 17, and 22 hybridize to M13mp98. Therefore, both DNA strands of the Bam HI insert of pPGA7 were cloned.

To confirm the cloning of the Bam HI insert of pPGA7, the recombinant M13mp8 and M13mp9 clones and an M13mp8 control were hybridized to ^{32}P-nick translated Bam HI insert of pPGA7. Figures 28a and b show the gel and autoradiogram for this experiment. All the M13 recombinants tested hybridized strongly to the labelled probe. Only the control M13mp8 did not hybridize. Clones M13mp8E5 (lane 12) and M13mp8E8 (lane 8) which contain complimentary sequences were chosen for dideoxy sequencing. The

FIGURE 26: DIGE test for the presence of M13 recombinants of a Bam HI digest of pPGA7

The viral DNAs were electrophoresed on 1% agarose. The following is a list of the lanes and the corresponding viral DNAs: 1, M13mp8 (control); 2, M13mp83; 3, M13mp84; 4, M13mp85; 5, 8.5 kb viral DNA (control); 6, 10.5 kb viral DNA (control); 7, M13mp97; 8, M13mp98; 9, M13mp99; 10, M13mp8 (control).

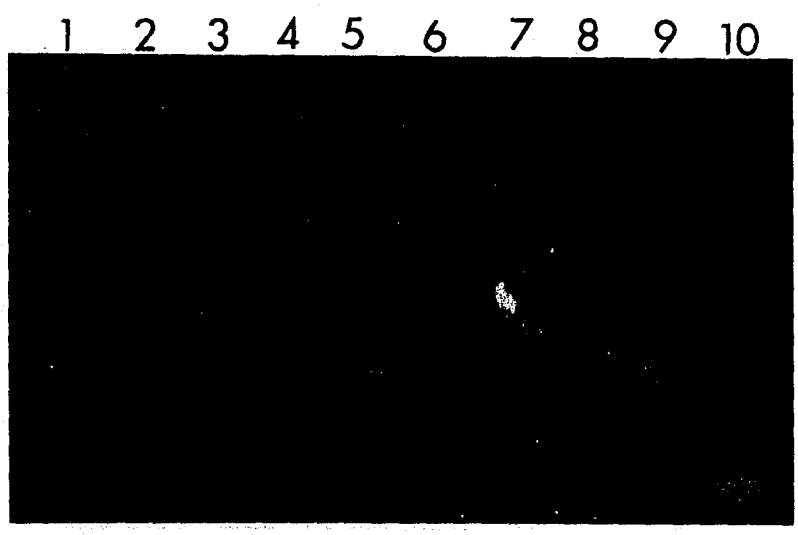
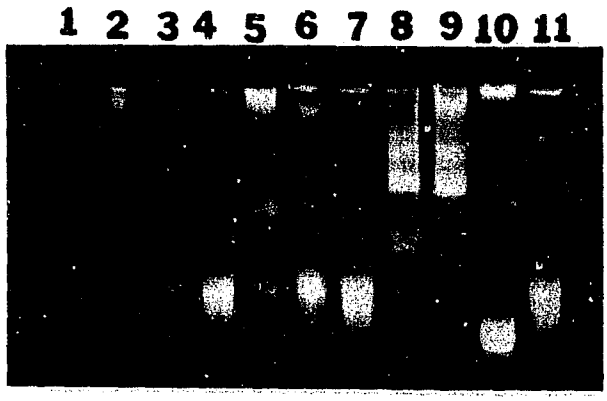

1 2 3 4 5 6 7 8 9 10

FIGURE 27: C-test of M13 recombinants containing the Bam H1 insert of pFGA7

(a) Those clones hybridized to M13mp85. The following is the list of the lanes and the corresponding clones: 1, M13mp85 plus M13mp8E1; 2, M13mp85 plus M13mp8E2; 3, M13mp85 plus M13mp8E4; 4, M13mp85 (control); 5, M13mp85 plus M13mp8E5; 6, M13mp85 plus M13mp8E6; 7, M13mp85 (control); 8, M13mp85 plus M13mp8E7; 9, M13mp85 plus M13mp8E8, 10, M13mp85 plus M13mp8 (control); 11, M13mp85 plus M13mp8E3.

(b) Those clones hybridized to M13mp98. The following is a list of the lanes and corresponding clones: 1, M13mp98 plus M13mp9G1; 2, M13mp98 plus M13mp9G3; 3, M13mp98 plus M13mp9G4; 4, M13mp98 (control); 5, M13mp98 plus M13mp9G5; 6, M13mp98 plus M13mp9G7; 7, M13mp98 (control); 8, M13mp98 plus M13mp9G8; 9, M13mp98 plus M13mp9G9; 10, M13mp98 plus M13mp8 (control); 11, M13mp98 plus M13mp99.

a



b

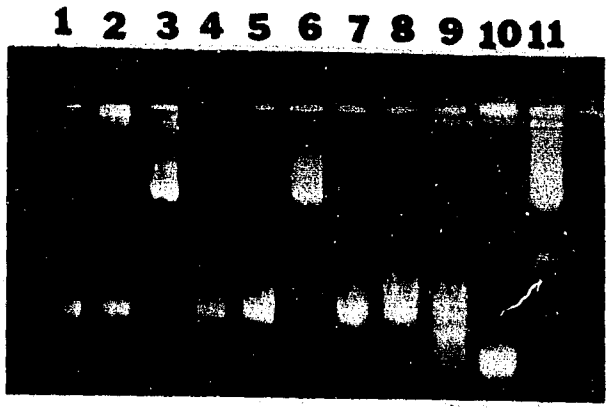
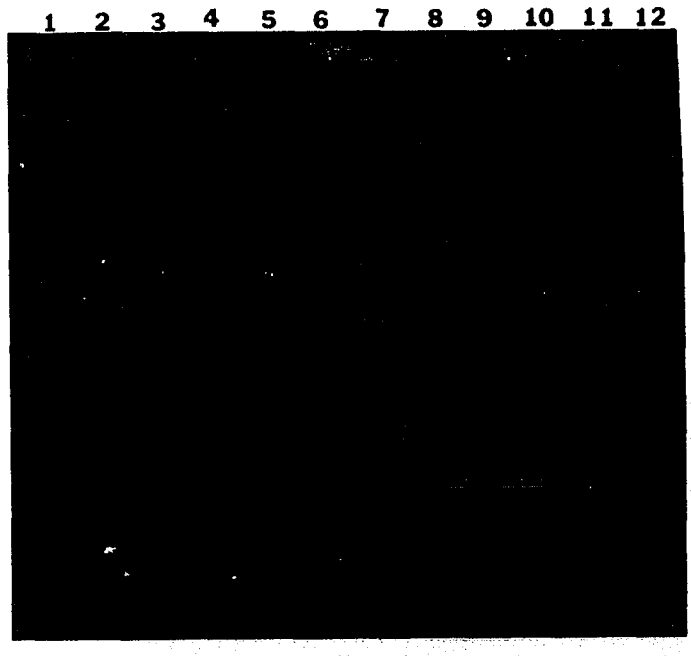


FIGURE 28: Gel and autoradiogram of M13 recombinants

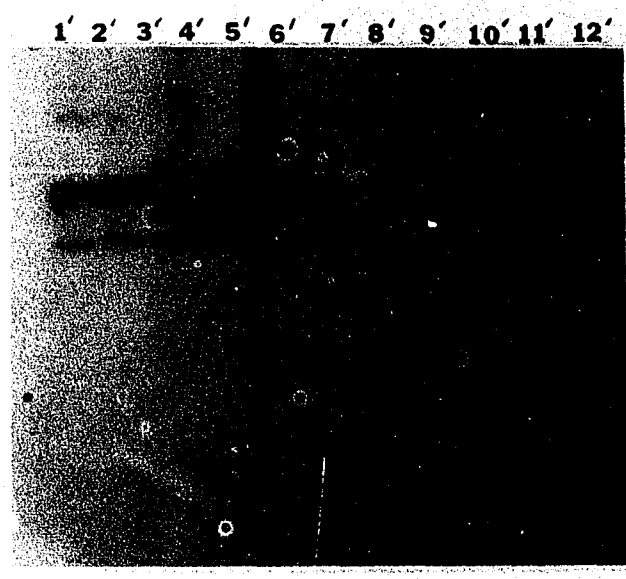
(a) B-test of M13 recombinants containing the Bam HI insert of pPGA7. The following is a list of the lanes and the corresponding viral clone DNAs: 1, M13mp9G8; 2, M13mp9G7, 3, M13mp9G3; 4, M13mp9G3; 5, M13mp99; 6, M13mp98; 7, M13mp8 (control); 8, M13mp8E8; 9, M13mp8E7; 10, M13mp8E2; 11, M13mp8E5; 12, M13mp85.

(b) Autoradiogram of the gel in 'a'. Nick transalted Bam HI insert of pPGA7 was used as the hybridization probe. The primed lane numbers in the autoradiogram correspond the the unprimed lane numbers in 'a'. All clones hybridize except the M13mp8 (lane 7) control.

a



b



sequencing gel is shown in Figure 33, and the DNA sequence read from this gel is shown in Table 5.

Subcloning and sequencing of the Hinc II-Eco RI fragments of pPGA7

pPGA7 was digested with Hinc II plus Eco RI and subcloned into the Hinc II and Eco RI sites in the RF forms of M13mp8 and M13mp9. By subcloning into both vectors, the DNA sequence could be read from both the Hinc II and from the Eco RI site towards the center of the cloned piece of DNA (Figure 29).

The double stranded replicative forms of some of these subclones were digested with Hinc II and Eco RI to determine the sizes of the pPGA7 Hinc II-Eco RI fragments which they contained. These gels are shown in Figures 30a and 30b. Clones M13mp8366, M13mp83611, M13mp83613, M13mp9404, M13mp94017, and M13mp94018 were all chosen for dideoxy sequencing. A dideoxy gel for M13mp83613 is shown in Figure 33.

The sequences obtained by dideoxy sequencing are shown in Table 5. The location of these sequences in pPGA7 is shown in Figure 31a. Clones M13mp8366 and M13mp83613 should be the same, but their sequences are different. The actual sequence of clone M13mp8366 is unusual in two respects. The sequence should start at a Hinc II site, but it does not; and part of the viral sequence, including the viral Hinc II site, has been deleted. The events which led to this cloning event are not understood, but cast doubt on the location of the M13mp8366 sequence. The insert in clone M13mp83611 was measured to be 0.7 kb. There are two possible partial digestion fragments of pPGA7 which could result in the 0.7 kb insert of M13mp83611; the Eco RI site in pBR322 to the Eco RI site of the

FIGURE 29: A schematic representation of the cloning of a Hinc II-Eco RI fragment into the vectors M13mp8 and M13mp9

The upper strand (5′-3′) of each vector is the viral strand which is the template for dideoxy sequencing. The region where the sequencing primer DNA binds and the direction of sequencing is indicated above each vector.

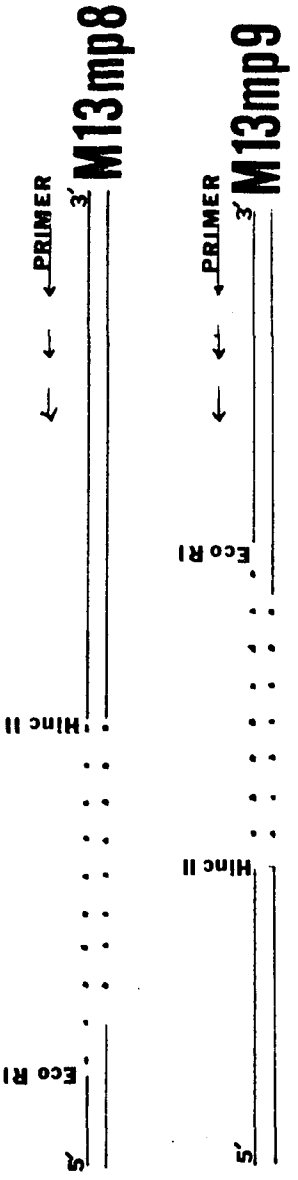

M13mp8
PRIMER
3'
Hinc II
Eco RI
5'
M13mp9
PRIMER
3'
Eco RI
Hinc II
5'

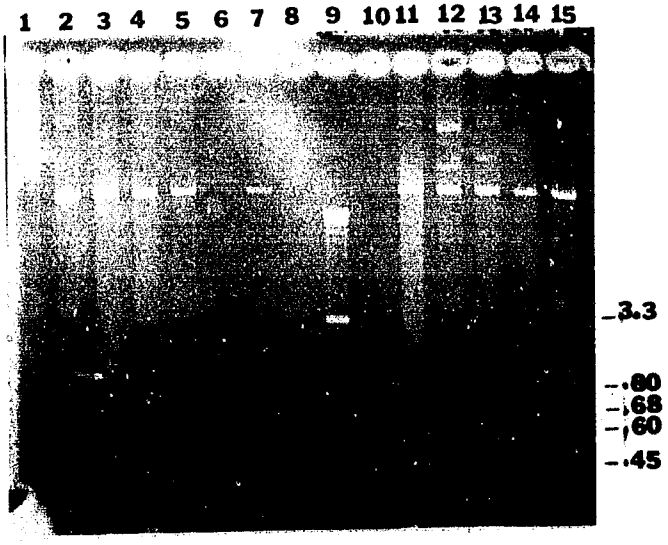
FIGURE 30: Hinc II plus Eco RI digests of M13 recombinants

(a) Hinc II plus Eco RI digestion of the M13mp8 recombinants of Hinc II plus Eco RI digested pPGA7. The following is a list of the lane numbers and the corresponding clones (numbers in parentheses are the length of the cloned pPGA7 Hinc II plus Eco RI fragment): 1, M13mp8362 (uncut); 2, M13mp8363 (0.6 kb); 3, M13mp8364 (1.3 kb and 0.6 kb); 4, M13mp8365 (0.6 kb); 5, M13mp8366 (0.6 kb); 6, M13mp8361 (uncut); 7, Hinc II plus Eco RI digest of pPGA7 (control); 8, M13mp8367 (?); 9, M13mp8368 (0.67 kb); 10, M13mp8369 (0.6 kb); 11, M13mp83611 (1.3 kb); 12, M13mp83613 (0.6 kb).

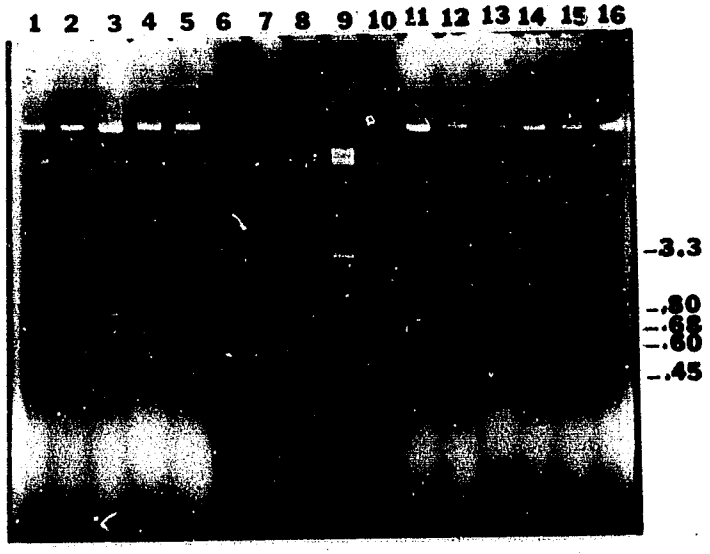
(b) Hinc II plus Eco RI digestion of the M13mp9 recombinants of Hinc II plus Eco RI digested pPGA7. The following is a list of the lane numbers and the corresponding clones (numbers in parentheses are the length of the cloned pPGA7 Hinc II plus Eco RI fragment): 1, M13mp9401 (0.67 kb); 2, M13mp9402 (0.67 kb); 3, M13mp9404 (0.8 kb); 4, M13mp9406 (0.67 kb); 5, M13mp9407 (0.45 kb); 6, M13mp9401 (uncut); 7, Hinc II plus Eco RI digest of pPGA7 (control); 8, M13mp9408 (0.8 kb); 9, M13mp94010 (0.6 kb); 10, M13mp94015 (0.6 kb); 11, M13mp94018 (0.6 kb); 12, M13mp94019 (0.6 kb); 13, M13mp94020 (0.67 kb).

Lengths shown are in kilobases.

a


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
3.3
.60
.68
.60
.45

b


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
3.3
.80
.68
.60
.45

M13 subclone, or the Hinc II site in the insert of pPGA7 to the first Hinc II site in pBR322. The sequencing gel for M13mp83611 was run too long, so the beginning restriction site is unknown. Because of these uncertainties, the sequence cannot be located on the map in Figure 31a.

There is a partial sequence overlap between clones M13mp8366 and M13mp94018 suggesting internal repetition in the pPGA7 clone.

Sequencing of other fragments of pPGA7

pPGA7 was digested with Hae III and Hpa II, and an attempt was made to subclone the resultant fragments into M13mp8. Clones were obtained and dideoxy sequencing analysis performed. The sequences obtained for 15 out of 26 clones were either M13mp8 or M13mp8366. The reappearance of the M13mp8366 sequence is not understood and therefore it was decided that the subcloning experiment was not successful. The 0.25 kb Hpa II-Hpa II fragment was electroeluted from acrylamide and taken to Dr. John Donelson's laboratory in Iowa City for Maxam-Gilbert sequencing. The fragment was cleaved with Hae III and end labelled. Because the Hae III site is very near one end of the fragment, only the sequence from left to right (see Figure 31a) could be read. This sequence is in Table 5, and the autoradiogram of the gel is shown in Figure 34. Its location on the map of pPGA7 is shown in Figure 31a.

Subcloning and sequencing of the 0.25 kb Pst I fragment of pPGB9

The 0.25 kb Pst I fragment of clone pPGB9 was electroeluted from 7% acrylamide and subcloned into M13mp8. Two resultant subclones were sequenced by the dideoxy method.

Figure 32 shows a Pst I digest of these subclones. In all cases, the

FIGURE 31: Locations of the subclones of pPGA7 and pPGB9 on the restriction endonuclease maps of pPGA7 and pPGB9

(a) A representation of the M13 subclones of pPGA7 and where their location is on the pPGA7 restriction map. The line above the map indicates the location of the repetitive sequence. The arrows below the map indicate the clones, the direction of sequencing, and the approximate distance sequenced. The numbers of the arrows are listed in parentheses before each clone. The lengths of the inserts of the various clones sequenced are listed below:

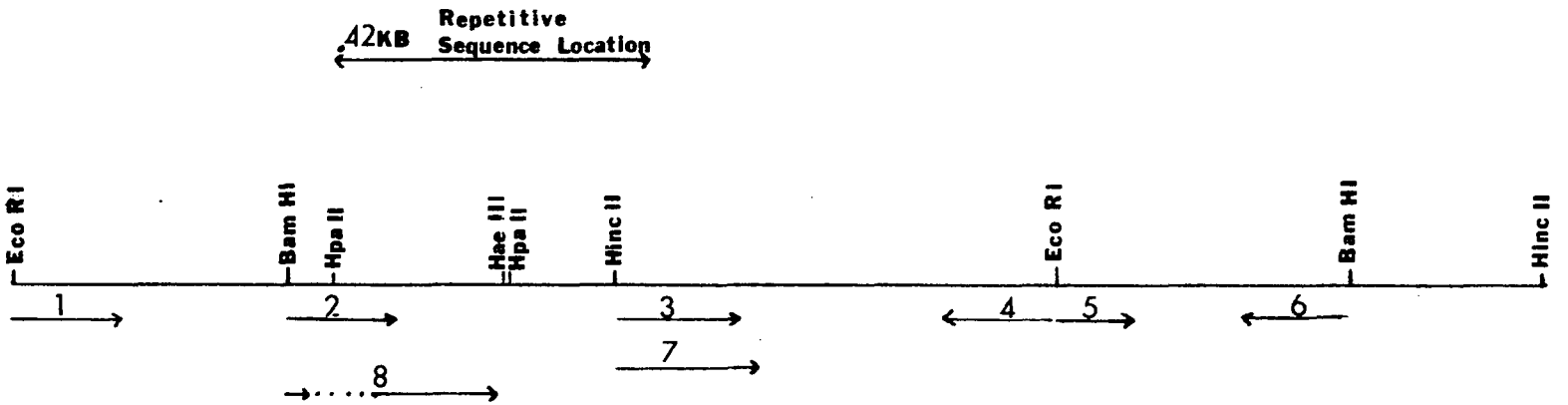
(6) M13mp8E5	1.40 kb	(1) M13mp9404	0.80 kb	(5) M13mp94017	0.67 kb
(2) M13mp8E8	1.40 kb	(4) M13mp94018	0.60 kb	(3) M13mp83613	0.06 kb
(7) M13mp8366	0.60 kb				

M13mp83611 1.3 kb (cannot be definitely located)

(8)Hpa II fragment 0.25 kb

(b) A representation of the M13 subclones of pPGB9 and where their location is on the pPGB9 restriction map. The line above the map indicates the location of the repetitive sequence. The arrows below the map indicate the clones, the direction of sequencing, and the approximate distance sequenced. The numbers of the arrows are listed in parentheses before each clone. The lengths of the inserts of the two clones sequenced are listed below:

(1 or 2) M13mp8PB9H or M13mp8PB9H 0.24 kb


42KB
Repetitive
Sequence Location
Eco RI
Bam HI
Hpa II
Hae III
Hpa II
Hinc II
Eco RI
Bam HI
Hinc II
1
2
3
4
5
6
7
8

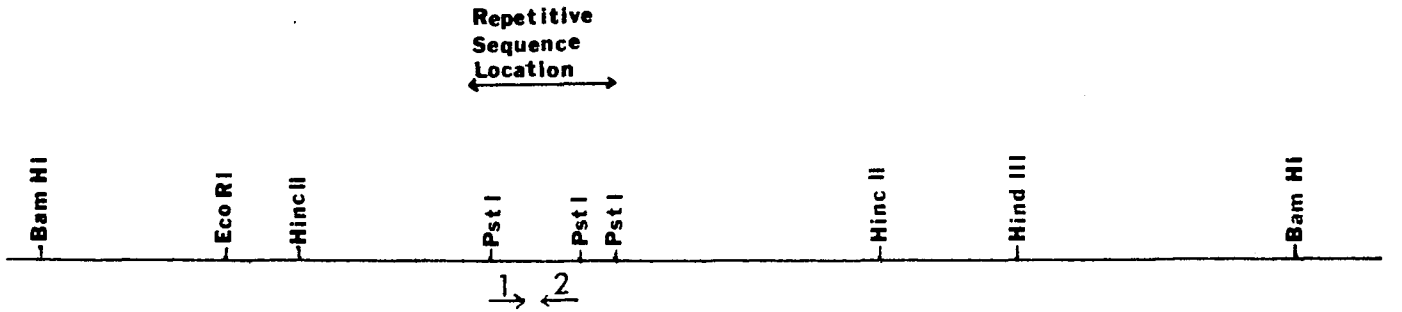

Repetitive
Sequence
Location
Bam HI
Eco RI
HincII
Pst I
Pst I
Pst I
Hinc II
Hind III
Bam HI
1
2

FIGURE 32: Pst I digest of the M13 subclones of the 0.25 Kb Pst I fragment of pPGB9

The following is a list of the lanes and the corresponding clones: 1, M13mp8 plus Hae III used as a length standard. Lengths shown are in kilobases. 2, M13mp8PB9F; 3, M13mp8PB9G; 4, M13mp8PB9H; 5, M13mp8PB9I; 6, M13mp8PB9J; 7, M13mp8PB9K; 8, M13mp8PB9L.

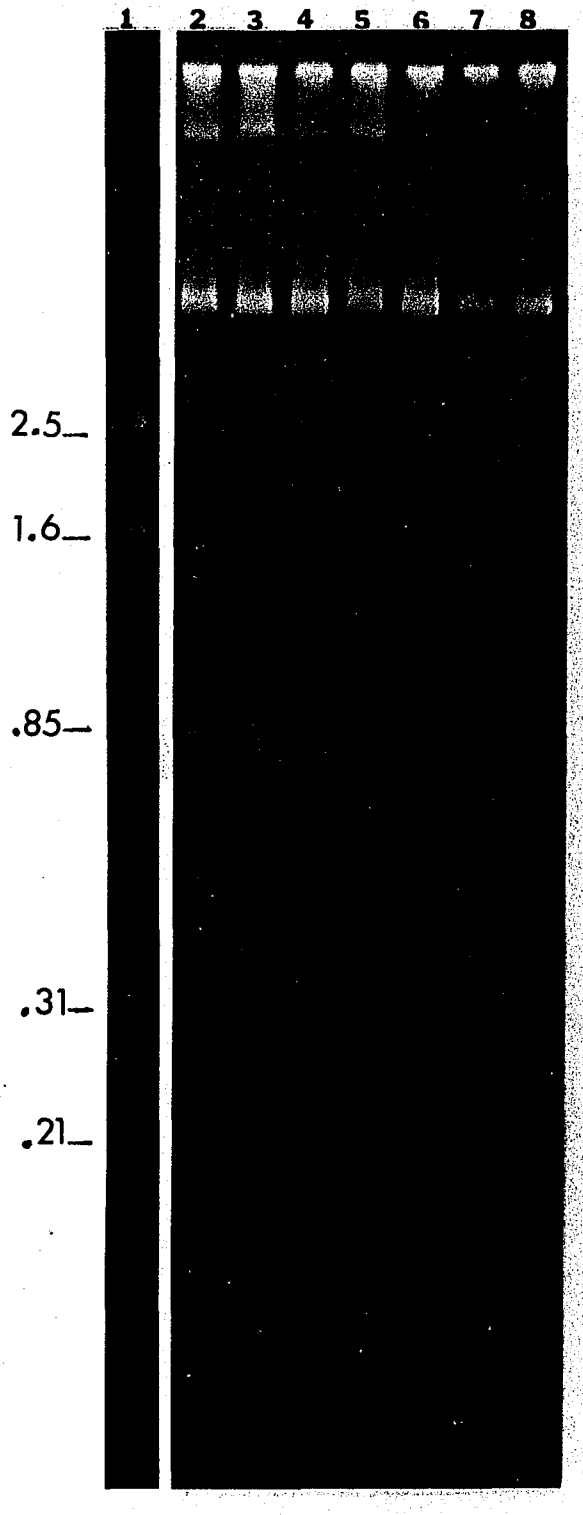

1
2
3
4
5
6
7
8
2.5
1.6
.85
.31
.21

FIGURE 33: Dideoxy DNA sequencing gel

2 hour and 4 hour electrophoresis of clones M13mp8E5, M13mp8E8 and M13mp83013. Four sequencing reactions are performed for each clone. They are denoted G, A, T, C, and the lanes are in this order from left to right. The sequence is read in the 5′ to 3′ direction starting at the bottom of the gel. The sequence shown in the 4 hour gel is the same as that bracketed in the 2 hour gel.


E5
E8
3013
G A T C G A T C G A T C
E5
E8
3013
G A T C G A T C G A T C
2 HRS.
4 HRS.

FIGURE 34: Maxam-Gilbert DNA sequencing gel and a part of the sequence for the 0.25 kb HapII fragment of pPGA7

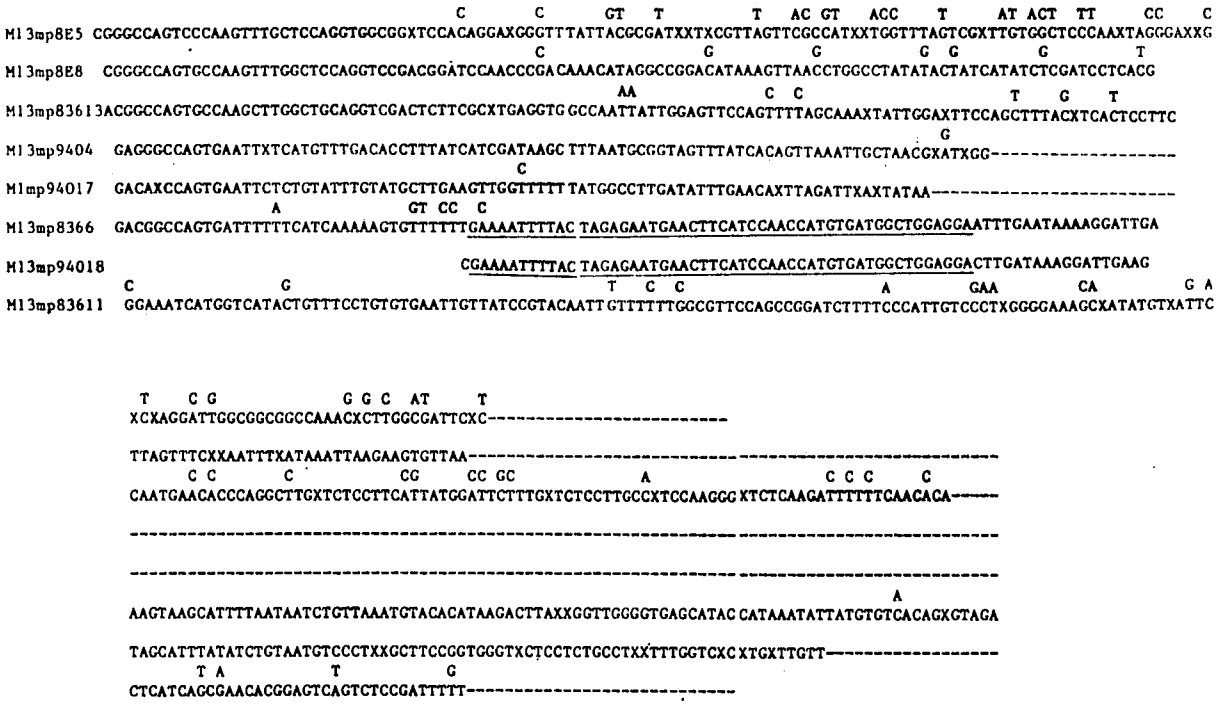
Four sequencing reactions are performed for each clone. The are denoted A, G, T, C, and are in this order from left to right. The sequence is read in the 5′ to 3′ direction starting at the bottom of the gel. A sample sequence reading is shown in the picture.


A G C T
3'
C
T
G
A
G
C
G
5'

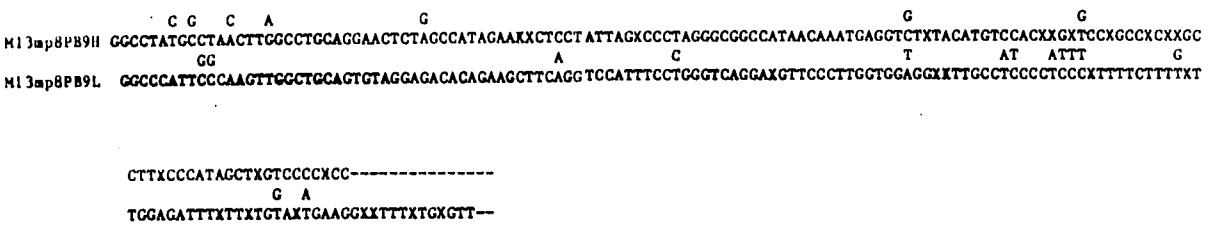
TABLE 5: DNA sequences determined for the various clones listed

The letters above the sequences are those for which two nucleotides could be read. The 'X' is where that particular base could not be determined. The boldface type is the M13 vector sequences, the underlining is the sequence homology between clones M13mp8366 and M13mp94018, and the symbol (....*) is the 5 nucleotide repeat in the Maxam-Gilbert sequences.

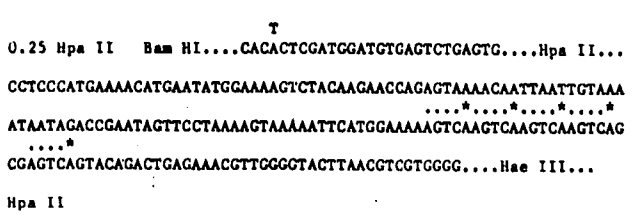
Dideoxy Sequencing of pPGA7



Dideoxy Sequencing of pPGB9



Maxam-Gilbert Sequencing of pPGA7



subcloned piece of DNA is the same length as the 0.25 kb Pst I fragment of pPGB9. Figure 31b shows the location of the sequences obtained. The two clones sequenced were M13mp8B9PH and M13mp8B9PL (Figure 32, lanes 4 and 8). The sequences for these fragments are listed in Table 5.

DISCUSSION

The experiments described in this thesis have partially characterized a family of short interspersed repeated DNA sequences found in the bovine genome. The restriction endonuclease maps (Figures 16a and b) of two of the recombinant plasmids, pPGA7 and pPGB9, show that the portion of each clone made up of repetitive DNA is limited to a few hundred nucleotides in length.

Figures 19-21 and Table 3 show that most of the recombinant plasmids containing repetitive bovine DNA, cross hybridize to the Bam HI inserts of either pPGA7 or pPGB9. The bovine recombinants represent random segments of the bovine DNA, selected only for the presence of repetitive sequences; therefore, the cross hybridization experiments suggest that there is only one major short interspersed repeated DNA sequence family in the bovine genome. In these same experiments, it is seen that the bovine repeated sequence represented in pPGA7 does not hybridize with the human Alu clone, BLUR-8 (Figure 20, lanes 23 and 23′). Table 3 also shows that there is no sequence in the bovine genome which hybridizes with the human Alu clone, BLUR-8. This implies that there is no Alu-like sequence in bovine DNA. This is supported by the DNA sequencing data from clones pPGA7 and pPGB9.

In other hybridization experiments, the repetitive DNA sequence in pPGA7 and pPGB9 is shown to comprise about 6% of the bovine genome (Table 4). The complexity of the bovine genome is 3×10^9 bp (Britten, 1971). If it is assumed that the short interspersed repeated sequences are about 300 nucleotides in length, then the repeated sequence family represented by

pPGA7 is present in about 600,000 copies in the bovine genome. Watanabe (et al., 1982) have discovered a short repetitive DNA sequence in the bovine genome which seems to be repeated about 100,000 times. The partial DNA sequences of pPGA7 and pPGB9 do not seem to be homologous to this reported sequence.

It is known that 25-30% of the bovine genome is made up of repetitive sequences (Britten and Smith, 1971; Mayfield, et al., 1980). About two-thirds of these repeated sequences, or 20% of the genome, is made up of the highly repetitious satellites. The other third, or 6-10% of the genome, is composed of the short interspersed repeated sequences (Mayfield et al., 1980). Since the hybridization experiments (Table 4) have shown that about 6% of the bovine genome is composed of repeated sequences which cross hybridize with pPGA7, and since 22 out of 26 repetitive sequence containing clones belong to this family, it is unlikely that there is any other major family of short interspersed repeated sequences in the bovine genome. As discussed in the introduction, a single major short interspersed repeated sequence family has been documented to occur in several other higher eukaryotic genomes.

The human Alu family was the first major class of short interspersed repeated sequences discovered. From hybridization experiments, as well as DNA sequencing data, the major family of short interspersed repeated sequences in the bovine genome is not a member of the Alu family. This does not mean that the function of the bovine family is different from that of the Alu family.

Like the human Alu family, the bovine sequence is probably a dimer or

multimer. The sizing experiments (Figures 24 and 25) indicate that the basic unit of the repeat in bovine DNA is about 120-130 nucleotides in length. This is determined by the finding of an S1 nuclease-resistant fraction in experiments where only tracer amounts of radioactive Bam HI inserts of pPGA7 and pPGB9 are allowed to self-hybridize (Figures 24 and 25). If there were no internal homology in the cloned DNA, then, at the concentrations of tracer used, there should be no S1-resistant molecules in hybridization experiments containing only tracer DNA. These results contrast with an average length of 350 bp by electronmicroscopy (Mayfield, et al., 1880) and 300, 400, and 600 bp as determined by S1 nuclease digestion of total calf thymus DNA (Figure 23). A dimeric structure is found in the human and monkey Alu families (Jelinek and Schmid, 1982; Schmid and Jelinek, 1982). The Alu-like sequence families of the Chinese hamster, mouse and rat are closely related to the left half of the human Alu dimer.

There has been speculation, and indeed some experimental evidence that the Alu family of short interspersed repeated sequences is an active, or at least a remnant of a transposable element. One feature of transposable elements is the presence of short direct repeats at the 3′ and 5′ ends. In the Maxam-Gilbert sequence data of the 0.25 kb Hpa II-Hpa II fragment of pPGA7 (Table 5), there is a short direct repeat of the sequence AGTCT. This sequence is repeated 5 times. The map position of this fragment is consistent with the possible 3′ start of the bovine repeated sequence. Even though the bovine sequence does not seem to be closely sequence related to the Alu sequence, there is a short region of

homology to a portion of the Alu sequence and monkey spcDNA (Krolewski, et al., 1982). This is the sequence, AGAGAATCA, which is found in the bovine clone, M13mp8366 (Table 5).

From the bovine sequence available, there is no sequence which compares with that found in RNA polymerase transcripts of the Alu family in humans (Duncan, et al., 1981), or that found near the origin of replication of SV40 or polyoma virus (McCutchan, and Singer, 1981; Deininger, et al., 1980). It is also not known whether the bovine short interspersed repeated sequence is present as snRNAs hydrogen bonded to hnRNA, or present as a bovine 7S cytoplasmic RNA.

In summary, these experiments have shown that the bovine short interspersed repeated sequence family found in clones pPGA7 and pPGB9;

1) makes up about 6% of the bovine genome

2) includes approximately 600,000 copies in the bovine genome

3) seems to have an internal region of homology which suggests a dimer or multimer composed of 120-130 bp units

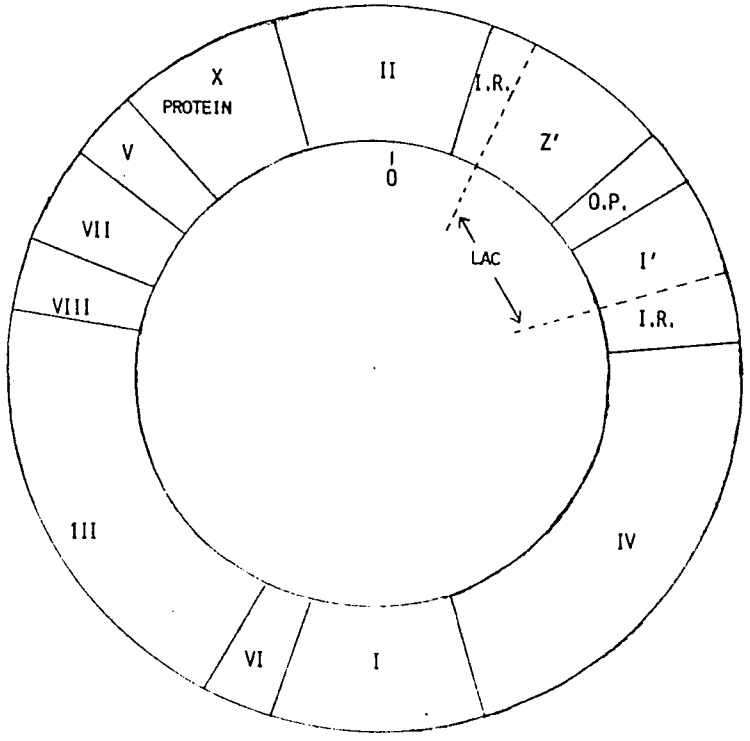
4) has 5 copies of a short, direct repeated sequence near the 3′ end

5) has no extensive homology with the Alu and Alu-like sequence families, but does have a 9 bp sequence homologous to a sequence found in the human Alu and monkey spcDNA sequences

6) has no extensive sequence homology with the bovine repeated sequences discovered by Watanabe (et al., 1982).

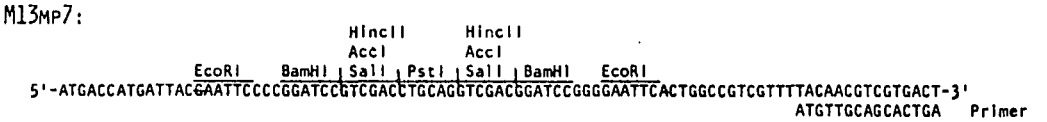
APPENDIX A

M13mp7, 8, or 9

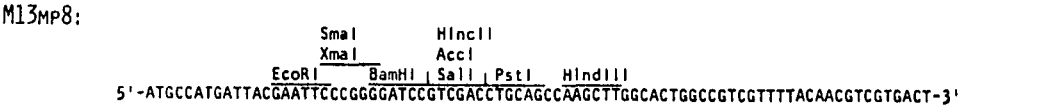


M13 Cloning Sites

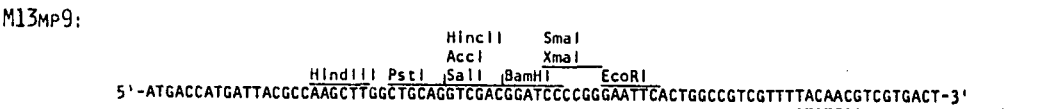
M13MP7:



M13MP8:



M13MP9:



APPENDIX B

Bacterial Growth Media and Plates

LB Medium: 10.0 grams bacto tryptone, 10.0 grams NaCl, 5.0 grams yeast extract/liter of water. Mix until dissolved and autoclave for 30 minutes.

LB Plates: 15.0 grams bacto agar/liter of LB medium. Autoclave for 30 minutes, let cool to 60^{o}C, then pour into petri plates. Yield: 30 plates/liter.

LB Ampicillin Plates: Make LB plates as above. Let media cool to 60^{o}C, then make it 200 ugs/ml ampicillin. Use an ampicillin stock solution of 20 mgs/ml in dimethyl sulfoxide (DMSO).

LB Tetracycline Plates: Make LB plates as above. Let cool to 60^{o}C, then make it 50 ugs/ml tetracycline. Use a tetracycline stock solution of 15 mgs/ml.

YT Medium: 10.0 grams bacto tryptone, 5.0 grams NaCl, 5.0 grams yeast extract/liter of water. Mix solution, then autoclave for 30 minutes.

YT Plates: 15.0 grams bacto agar/liter of YT medium. Autoclave for 30 minutes, let cool to 60^{o}C, then pour plates. Yield 25-30 plates/liter.

YT Soft Agar: 6.0 grams bacto agar/liter of YT medium. Autoclave for 30 minutes, then keep in a 42^{o}C water bath.

Minimal Plates: 15.0 grams bacto agar; 10.5 grams K_2HPO_4; 4.5 grams KH_2PO_4; 1.0 gram $(NH_4)_2SO_4$; 0.5 grams sodium citrate.H_2O; 0.2 grams

$MgSO_4.7H_2O$[a]; 5.0ugs thiamine-HCl[a]; 2.0 grams glucose[a]; per liter of water. The solutions with a superscript'a' were sterilized by autoclaving separately as 100-fold concentrated stock solutions, then added to the remaining sterilized solution when it had cooled to 45°C just prior to pouring the plates. Yield: 25-30 plates/liter.

Solutions and Buffers

Saline-EDTA: 70 mM NaCl, 30 mM EDTA (pH8)

TNKM: 0.05 M TRIS-HCl (pH6.7), 0.025 M KCl, 0.13 M NaCl, 3.0 mM $MgCl_2.6H_2O$.

TE: 10 mM TRIS, 1 mM EDTA (pH8).

KPB: equal molar amounts of monobasic and dibasic potassium phosphate

BPB Dye Mix: 0.1% bromophenol blue, 10% ficoll, 50% glycerol in water.

1x TBE: 89 mM TRIS, 89 mM boric acid, 3 mM EDTA (pH7.6).

1x TNE: 150 mM NaCl, 50 mM TRIS-HCl (pH7.5), 1 mM EDTA.

1x SSC: 0.15 M NaCl, 0.015 M sodium citrate.

1x Denhardts: 0.02%(w/v) ficoll (MW 400,000); 0.02%(w/v) polyvinylpyrollidone (MW 360,000); 0.2%(w/v) BSA; dissolved in 3x SSC.

Spermine Wash: 0.3 M sodium acetate, 0.1 M magnesium acetate, in 75% ethanol.

<u>Enzyme buffers</u>

Acc I: 6 mM TRIS-HCl (pH7.5), 6 mM NaCl, 6 mM $MgCl_2$, 100 ugs/ml BSA, 6 mM 2-mercaptoethanol.

Alu I: 6 mM TRIS-HCl (pH7.6), 50 mM NaCl, 6 mM $MgCl_2$, 6 mM 2-mercaptoethanol.

Bam HI: 20 mM TRIS-HCl (pH7), 100 mM NaCl, 7 mM $MgCl_2$, 2 mM 2-mercaptoethanol.

Hae III: 50 mM TRIS-HCl (pH7.5), 5 mM $MgCl_2$, 0.5 mM dithiothreitol.

Hinc II: 10 mM TRIS-HCl (pH7.9), 60 mM NaCl, 7 mM $MgCl_2$, 1 mM dithiothreitol.

Hpa II: 20 mM TRIS-HCl (pH7.4), 7 mM $MgCl_2$, 1 mM dithiothreitol.

Hinc II-Eco RI double digest: Hinc II buffer.

Bam HI-Hind III double digest: Bam HI buffer.

Low Buffer: 10 mM TRIS-HCl (pH7.4), 10 mM $MgSO_4$, 1 mM dithiothreitol.

Medium Buffer: 50 mM NaCl, 10 mM TRIS-HCl (pH7.4), 10 mM $MgSO_4$, 1 mM dithiothreitol.

High Buffer: 100 mM NaCl, 50 mM TRIS-HCl (pH7.4), 10 mM $MgSO_4$.

1x Nick Translation Buffer: 50 mM TRIS-HCl (pH7.5), 5 mM $MgCl_2$, 1 mM 2-mercaptoethanol.

1x S1 Nuclease Assay Buffer: 30 mM sodium acetate (pH4.6), 50 mM NaCl, 1 mM $ZnSO_4$, 5% glycerol.

<u>Buffers and nucleotides for dideoxy DNA sequencing</u>

Hin Buffer: 70 mM TRIS-HCl (pH7.5), 70 mM $MgCl_2$, 500 mM NaCl.

dGTP: 0.5 mM; dilute 10 mM dGTP stock solution 1:20 with water.

dATP: 0.5 mM; dilute 10 mM dATP stock solution 1:20 with water.

dTTP: 0.5 mM; dilute 10 mM dTTP stock solution 1:20 with water.

dCTP: 0.5 mM; dilute 10 mM dCTP stock solution 1:20 with water.

G': 20 uls dTTP, 20 uls dCTP, 1 ul dGTP, 20 uls Hin buffer.

A': 20 uls dTTP, 20 uls dCTP, 20 uls dGTP, 20 uls Hin buffer.

T': 1 ul dTTP, 20 uls dCTP, 20 uls dGTP, 20 uls Hin buffer.

C': 20 uls dTTP, 1 ul dCTP, 20 uls dGTP, 20 uls Hin buffer.

ddG: 0.5 mM; dilute 10 mM ddGTP stock solution 1:20 with water.

ddA: 0.25 mM; dilute 10 mM ddATP stock solution 1:40 with water.

ddT: 1 mM; dilute 10 mM ddTTP stock solution 1:10 with water.

ddC: 0.25 mM; dilute 10 mM ddCTP stock solution 1:40 with water.

MAKE FRESH NUCLEOTIDE SOLUTIONS EVERY TWO WEEKS

Formamide-dye Mix: 0.1%(w/v) xylene cyanol FF, 10 mM sodium EDTA, 0.1%(w/v) bromophenol blue, dissolved in formamide. This was made 20 mM NaOH just prior to use.

Buffers and solutions for Maxam-Gilbert DNA sequencing

XH Buffer: 10 mM TRIS-HCl (pH7.4), 150 mM NaCl, 6 mM $MgCl_2$, 6 mM 2-mercaptoethanol, 100 ugs/ml autoclaved gelatin.

L Buffer: 10 mM TRIS-HCl (pH7.4), 10 mM NaCl, 6 mM $MgCl_2$, 6 mM 2-mercaptoethanol, 100 ugs/ml autoclaved gelatin.

A Mix: 1.2 N NaOH, 1 mM EDTA (acid form).

C Mix: 3.75 M NaCl, 20 ugs/ml sheared, boiled salmon sperm DNA.

T Mix: 20 ugs/ml sheared, boiled salmon sperm DNA.

G Mix: 50 mM sodium cacodylate (pH8), 10 mM $MgCl_2$, 1 mM EDTA, 5 ugs/ml sonicated, boiled salmon sperm DNA.

A Stop: 1 M acetic acid.

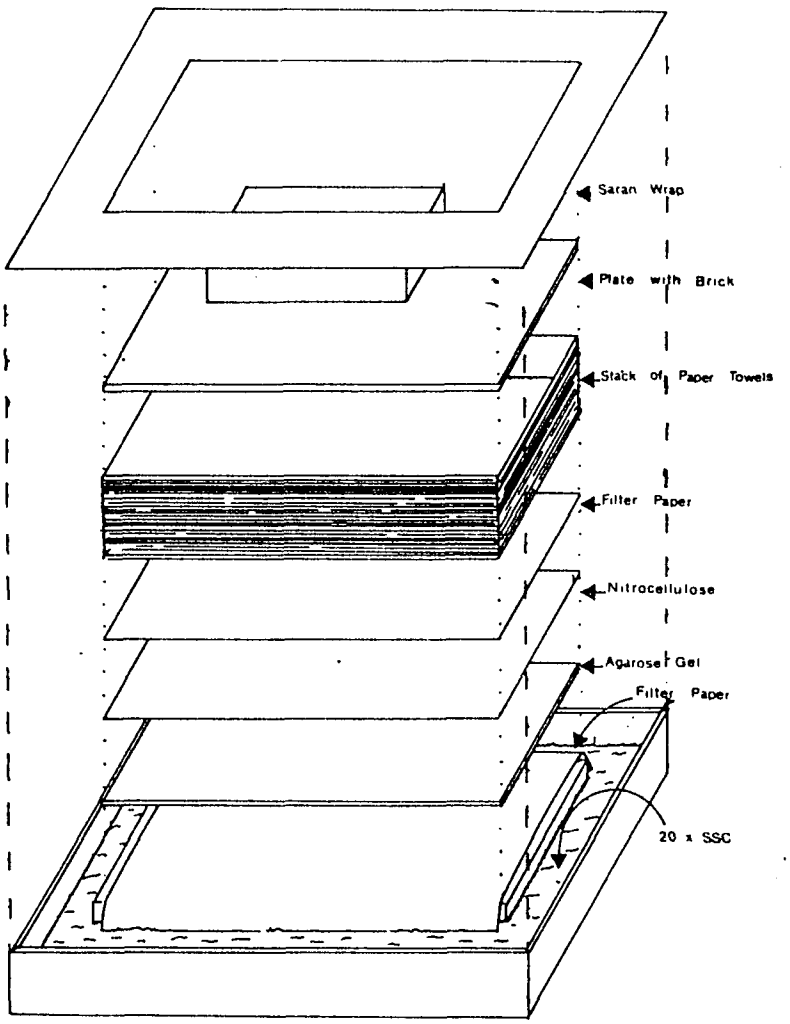
CT Stop: 0.3 M sodium acetate (pH7), 0.1 mM EDTA, 25 ugs/ml tRNA.

G Stop: 1.5 M sodium acetate (pH7), 100 ugs/ml tRNA

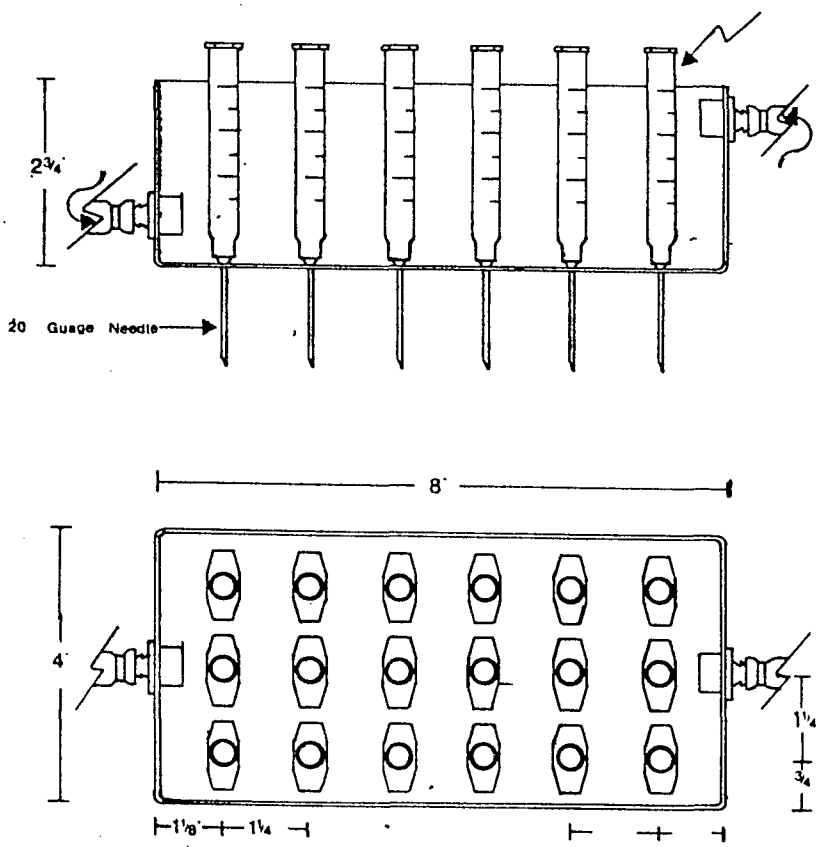
MG Dye Mix: 15 uls of 5 mM EDTA, 7.5 uls of 1 N NaOH, 75 uls of 0.5% xylene cyanol in 10 M urea, 52.5 uls water.

APPENDIX C

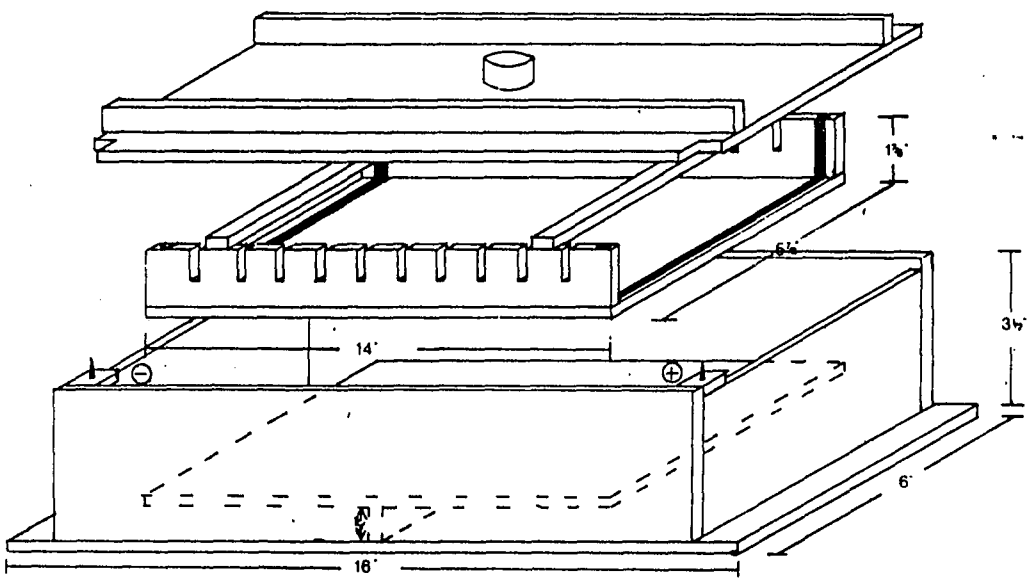
Southern Transfer



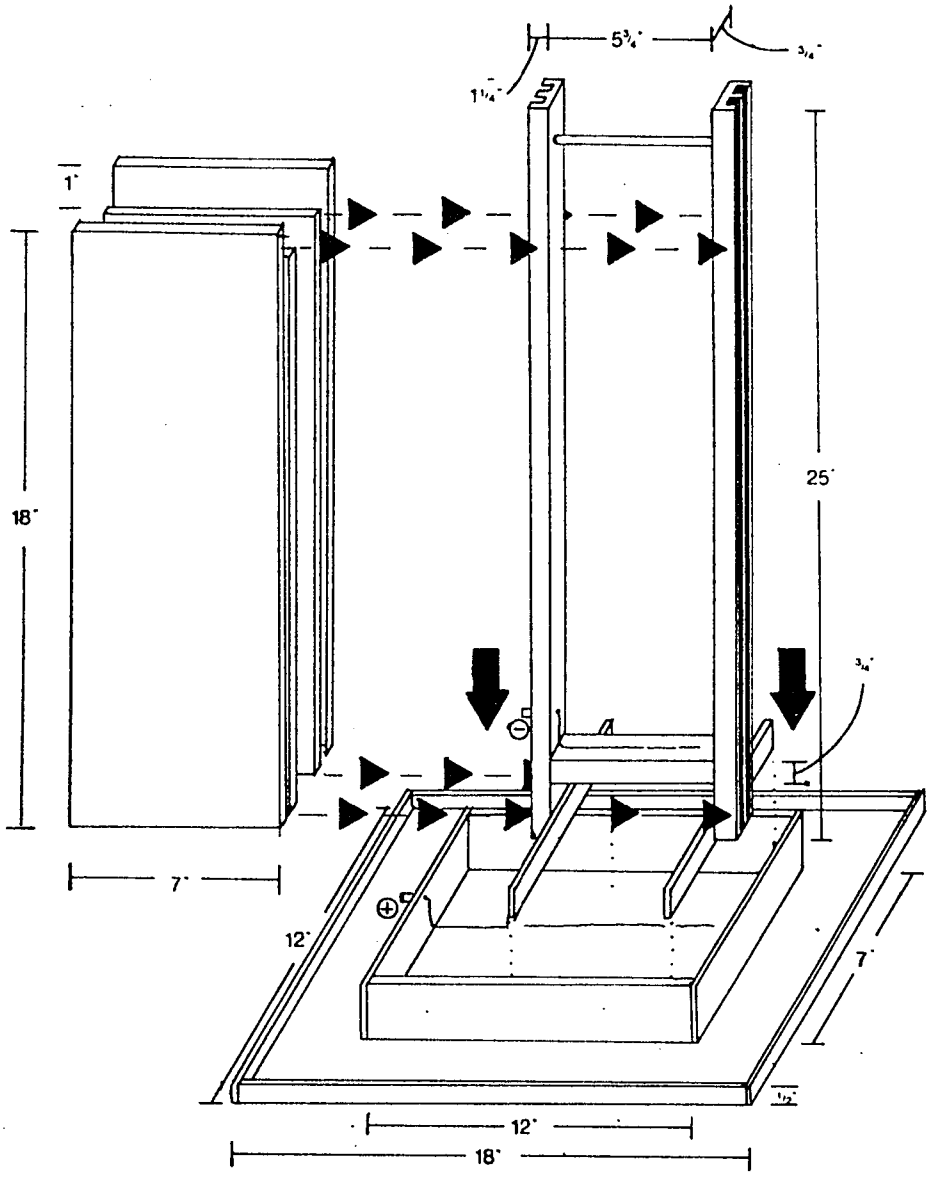
Hydroxylapatite Columns



Horizontal Electrophoresis Apparatus



DNA Sequencing Apparatus



APPENDIX D

Thermal Denaturation of DNAs using the S1 Nuclease Buffer

All DNAs have characteristic melting tamperatures, Tm, which in various neutral salts (Marmur and Doty, 1962) are related to the percent G-C content. The greater the G-C content, the higher the Tm. Since the S1 nuclease assay buffer is 0.01 M acetate (pH4.6) and $5x10^{-5}$ M $ZnSO_4$, it was decided to determine whether or not the same general relationship between G-C content and Tm held for this buffer.

To carry out these tests, a circulating water bath was connected to a jacketed four position cell holder in a Beckman spectrophotometer. The temperatures of the control and DNA-containing cuvettes were raised $1^{o}C$ every 3 minutes. A thermocouple was used to measure the temperature inside the control cuvette. Depending upon the experiment, the control cuvette contained either 0.01 M acetate (pH4.6) or 0.01 M acetate (pH4.6) plus $5x10^{-5}$ M $ZnSO_4$. The temperature and absorbance at 260nm for up to three DNA samples were recorded simultaneously from $40^{o}C$ to $75^{o}C$.

Figure 35 shows that $5x10^{-5}$ M $ZnSO_4$ raises the Tm of the calf thymus DNA by about $10^{o}C$ in 0.01 M acetate at pH4.6. To determine if zinc affected the general relationship between percent G-C and Tm, three DNAs of differing G-C content were simultaneously heat-denatured in the spectrophotometer. The three DNAs used were from Clostridium perfringens (31% G-C), E. coli (50% G-C), and calf thymus (43% G-C) (Sorber, 1968). Figure 36 is a graph of the temperature versus the Tm for these DNAs. This shows that the relationship of G-C content to Tm for these DNAs in

FIGURE 35: Thermal denaturation

Thermal denaturation of calf thymus DNA in 0.01 M sodium acetate (pH4.6) (.), and in 0.01 M sodium acetate (pH4.6) plus $5x10^{-5}$ M $ZnSO_4$ (x).

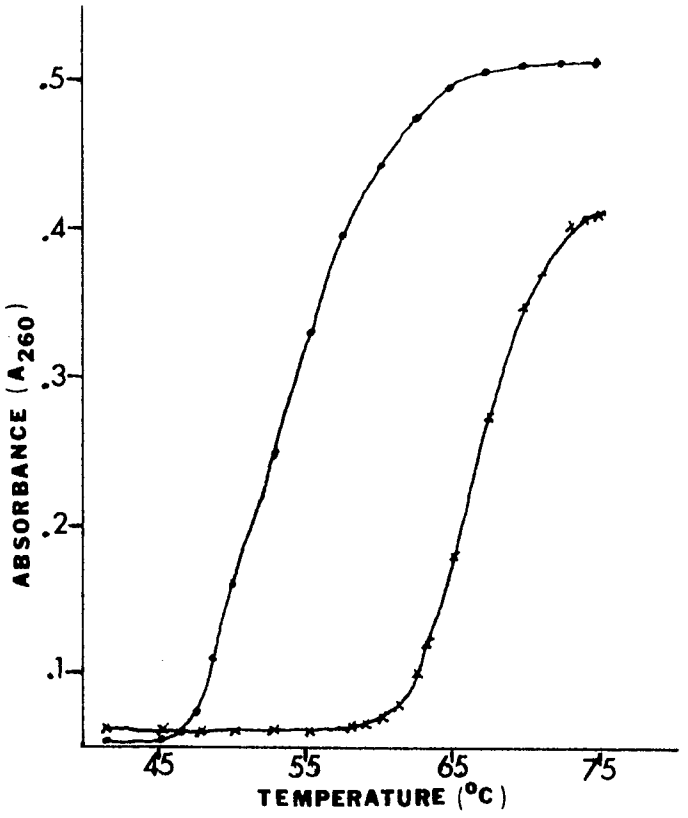
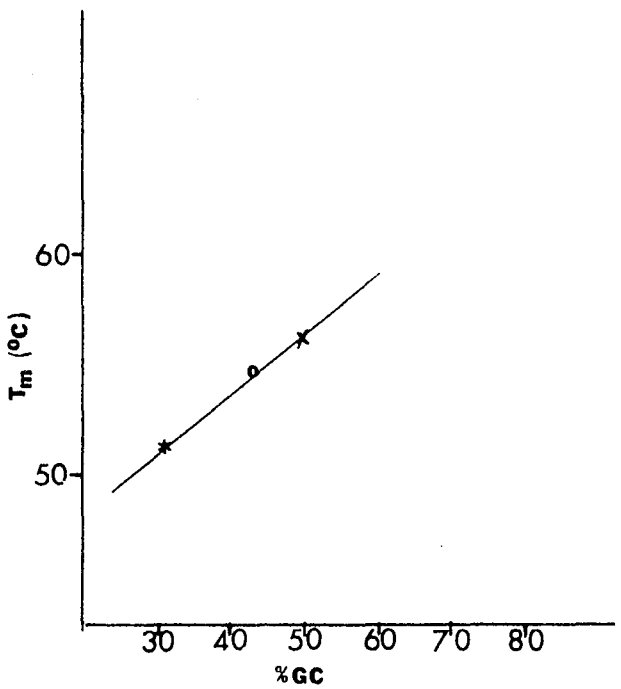

.5
.4
.3
.2
.1
ABSORBANCE (A_{260})
45
55
65
75
TEMPERATURE (°C)

FIGURE 36: Plot of Tm vs G-C content of three DNAs

Plot of the Tm for the three DNAs in the graph versus the percent G-C content of three DNAs in 0.01 M sodium acetate (pH4.6) plus $5x10^{-5}$ M $ZnSO_4$. (*) Clostridium perfringens DNA (31% G-C); (o), calf thymus DNA (43% G-C); (x), E. Coli DNA (50% G-C).


60
50
T_m (°C)
30
40
50
60
70
80
%GC

the presence of Zn^{++} ions at pH4.6 is similar to that observed at neutral pH in the absence of Zn^{++} ions.

BIBLIOGRAPHY

Adams, J. W., R. E. Kaufman, P. J. Kretschmer, M. Harrison, and A. W. Neinhuis. 1980. A family of long reiterated DNA sequences, one copy of which is next to the human beta globin gene. Nucleic Acids Res. 8:6113-6128.

Arrighi, Frances E., Manley Mandel, Janet Bergendahl, and T. C. Hsu. 1970. Buoyant densities of DNA of mammals. Biochem. Gen. 4:367-376.

Balmain, Alan, Robert Kromiauf, J. Keith Vass, and George D. Birnie. 1982. Cloning and characterisation of the abundant cytoplasmic 7S RNA from mouse cells. Nucleic Acids Res. 10:4259-4277.

Bethesda Research Laboratories. 1980. Instruction manual, M13 cloning/'dideoxy' sequencing. BRL Inc. (Bethesda Research Laboratories), Gaithersburg, Maryland. 49pp.

Birnboim, H. C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513-1523.

Bolivar, Francisco, Raymond Rodriguez, Mary C. Betlach, and Herbert W. Boyer. 1977a. Construction and characterization of new cloning vehicles: I ampicillin-resistant derivatives of the plasmid pMB9. Gene 2:75-93.

Bolivar, Francisco, Raymond L. Rodriguez, Patricia J. Greene, Mary C. Betlach, Herbert L. Heyneker, Herbert W. Boyer, Jorge H. Crosa, and Stanley Falkow. 1977b. Construction and characterization of new cloning vehicles: II a multipurpose cloning system. Gene 2:95-113.

Britten, Roy J. 1971. Observed properties of repeated DNA sequences. Carnegie Inst. of Wash. Year Book 68:376-378.

Britten, R. J., and E. H. Davidson. 1969. Gene regulation for higher cells: a theory. Science 165:529-540.

Britten, Roy J., and Eric H. Davidson. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origins of the evolutionary novelty. Q. Rev. Biol. 46:111-133.

Britten, Roy J., Dale E. Graham, and Berney R. Neufeld. 1974. Analysis of repeating DNA sequences by reassociation. Article 29 in Lawrence Grossman and Kivie Moldave eds. Methods in enzymology. Part E. Nucleic acids and protein synthesis. Academic Press, New York.

Britten, Roy J., Dale E. Graham, Francine C. Eden, Denise M. Painchaud, and Eric H. Davidson. 1976. Evolutionary divergence and length of repetitive sequences in sea urchin DNA. J. Mol. Biol. 9:1-23.

Britten, R. J. and D. E. Kohne. 1968. Repeated sequences in DNA. Science 161:529-540.

Britten, Roy J. and David E. Kohne. 1970. Repeated segments of DNA. Scientific American 222:24-31. (offprint #1173).

Britten, R. J. and Jean Smith. 1971. A bovine genome. Carnegie Inst. of Wash. Year Book 68:378-386.

Calabretta, Bruno, Donald L. Robberson, Abby L. Maizel, and Grady F. Sanders. 1981. mRNA in human cells contains sequences complimentary to the Alu family of repeated DNA. Proc. Natl. Acad. Sci. (USA) 78:6003-6007.

Calabretta, Bruno, Donald L. Robberson, Hugo A. Barrera-Saldana, Thymios P. Lambrou, and Grady F. Sanders. 1982. Genome instability in a region of human DNA enriched in Alu repeat sequences. Nature 296:219-225.

Cameron, John R., Elwyn Y. Loh, and Ronald W. Davis. 1979. Evidence for transposition of dispersed repetitive DNA families in yeast. Cell 16:739-751.

Clewell, Don B. 1972. Nature of ColE$_1$ plasmid replication in Escherichia coli in the presence of chloramphenicol. J. Bacteriol. 110:667-676.

Clewell, Don B. and Donald R. Helinski. 1970. Properties of a supercoiled deoxyribonucleic acid-protein complex and strand specificity of the relaxation event. Biochemistry 9:4428-4440.

Cohen, Stanley N., Annie C. Y. Chang, and Leslie Hsu. 1969. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. Proc. Natl. Acad. Sci. (USA) 69:2110-2114.

Davidson, Eric H. and Roy J. Britten. 1979. Regulation of gene expression: possible role of repetitive sequences. Science 204:1052-1059.

Davidson, Eric H., Barbara R. Hough, Christopher S. Amenson, and Roy J. Britten. 1973. General interspersion of repetitive with non-repetitive sequence elements in the DNA of Xenopus. J. Mol. Biol. 77:1-23.

Davidson, Eric H., Glen A. Galau, Robert C. Angerer, and Roy J. Britten. 1975. Comparative aspects of DNA organization in Metazoa. Chromosoma (Berl.) 51:253-259.

Deininger, Prescott L. and Carl W. Schmid. 1976. An electron microscope study of the DNA sequence organization of the human genome. J. Mol. Biol. 106:773-790.

Deininger, Prescott L. and Carl W. Schmid. 1979. A study of the evolution of repeated DNA sequences in primates and the existence of a new class of repetitive sequences in primates. J. Mol. Biol. 127:437-460.

Deininger, Prescott L., Abby Esty, Patricia LaPorte, Hingge Hsu, and Theodore Friedmann. 1980. The nucleotide sequence and restriction enzyme sites of the polyoma genome. Nucleic Acids Res. 8:855-860.

Deininger, Prescott L., Douglas J. Jolly, Carol M. Rubin, Theodore Friedmann, and Carl W. Schmid. 1981. Base sequence studies of 300 nucleotide renatured repeated human DNA clones. J. Mol. Biol. 151:17-33.

Doolittle, W. Ford and Carmen Sapienza. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284:601-603.

Duncan, Craig, H., Pundur Jagadeeswaran, Richard R. C. Wang, and Sherman M. Weissman. 1981. Structural analysis of templates and RNA polymerase III transcripts of Alu family sequences interspersed among the human B-like globin genes. Gene 13:185-196.

Flamm, W. G., P. M. B. Walker, and Mora McCallum. 1969. Some properties of the single strands isolated from the DNA of the nuclear satellite of the mouse (*Mus musculus*). J. Mol. Biol. 40:423-443.

Flavell, Andrew J. and David Ish-Horowicz. 1981. Extrachromosomal copies of the eukaryotic transposable element *copia* in *Drosophila* cells. Nature 292:591-595.

Grimaldi, Giovanna, Cary Queen, and Maxine F. Singer. 1981. Interspersed repeated sequences in the African green monkey genome that are homologous to the human alu family. Nucleic Acids Res. 9:5553-5568.

Gronenborn, Bruno, and Joachim Messing. 1977. Filamentous coliphage M13 as a cloning vehicle: insertion of a Hind II fragment of the *lac* regulatory region in M13 replicative form *in vitro*. Proc. Natl. Acad. Sci. (USA) 74:3642.

Gronenborn, Bruno, and Joachim Messing. 1978. Methylation of single-stranded DNA *in vitro* introduces new restriction endonuclease cleavage sites. Nature 272:375-377.

Grossman, Lawrence and Kivie Moldave eds. 1980. Methods in enzymology. Volume 65. Nucleic acids. Part I. Academic Press, New York. 968pp.

Harada, Fumio and Nobyyuki Kato. 1980. Nucleotide sequences of the 4S RNAs associated with poly(A)-containing RNAs of mouse and hamster cells. Nucleic Acids Res. 8:1273-1285.

Haynes, Susan R. and Warren R. Jelinek. 1981. Low molecular weight RNAs transcribed *in vitro* by RNA polymerase III from Alu-type dispersed repeats in Chinese hamster DNA are also found *in vivo*. Proc. Natl. Acad. Sci. (USA) 78:6130-6134.

Haynes, Susan R., Thomas P. Toomey, Leslie Leinwand, and Warren R. Jelinek. 1981. The Chinese hamster alu-equivalent sequence: a conserved, highly repetitious, interspersed deoxyribonucleic acid sequence in mammals has a structure suggestive of a transposable element. Mol. Cell. Biol. 1:573-583.

Houck, Catherine M., Frank P. Rinehart, and Carl W. Schmid. 1978. Fractionation of renatured repetitive human DNA according to thermal stability, sequence length, and renaturation rate. Biochim. Biophys. Acta 518:37-52.

Houck, Catherine M., Frank P. Rinehart, and Carl W. Schmid. 1979. A ubiquitous family of repeated DNA sequences in the human genome. J. Mol. Biol. 132:289-306.

Hoyer, B. H., B. J. McCarthy, and E. T. Bolton. 1964. A molecular approach to the systematics of higher organisms. Science 144;959-967.

Jacobsen, H., H. Klenow, and K. Ovargaard-Hansen. 1974. The N-terminal amino acid sequences of DNA polymerase I from *Escherichia coli* and of the large and small fragments obtained by a limited proteolysis. Eur. J. Biochem. 45:623.

Jelinek, Warren R. and Carl W. Schmid. 1982. Repetitive sequences in eukaryotic DNA and their expression. Ann. Rev. Biochem. 51:813-844.

Jelinek, Warren R., Thomas P. Toomey, Leslie Leinwand, Craig H. Duncan, Paul A. Biro, Prabhakrar V. Choudary, Sherman M. Weissman, Carol M. Rubin, Catherine M. Houck, Prescott L. Deininger, and Carl W. Schmid. 1980. Ubiquitous, interspersed repeated sequences in mammalian genomes. Proc. Natl. Acad. Sci. (USA) 77:1398-1402.

Kohne, D. E. and R. Britten. 1975. Hydroxyapatite techniques for nucleic acid reassociation. Nucleic Acids Res. 2:500-512.

Krolewski, John J., Arthur H. Bertelsen, M. Zafri Humayun, and Mark G. Rush. 1982. Members of the Alu family of interspersed, repetitive DNA sequences are in the small circular DNA population of monkey cells grown in culture. J. Mol. Biol. 154:399-415.

Lewin, Benjamin. 1977. Gene expression 3, plasmids and phages. John Wiley & Sons, New York. 925pp.

Lewin, Benjamin. 1980. Gene expression 2. Second edition. Eukaryotic chromosomes. John Wiley & Sons, New York. 1160pp.

Lewin, Roger. 1982. Surprising discovery with a small RNA. Science 218:777-778.

Mandel, M. and A. Higa. 1970. Calcium-dependent bacteriophage DNA infection. J. Mol. Biol. 153:159-162.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning/ a laboratory manual. Cold Spring Harbor, New York 545pp.

Manning, Jerry E., Carl W. Schmid, and Norman Davidson. 1975. Interspersion of repetitive and nonrepetitive DNA sequences in the Drosophila melanogaster genome. Cell 4:141-155.

Marmur, J. and P. Doty. 1962. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. J. Mol. Biol. 5:109-118.

Marx, Kenneth A. 1980. Non-satellite repetitive human DNA families sequence properties and evidence for occurrance in chimpanzee DNA. Biochim. Biophys. Acta 608:232-242.

Maxam, A. M., W. Gilbert. 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. (USA) 74:560.

Mayfield, John E. and John R. Ellison. 1976. The fluorescence localization of mouse satellite DNA in interphase nuclei. Chromosoma (Berl.) 54:27-31.

Mayfield, John E., Julie F. McKenna, and Bruce S. Lessa. 1980. The sequence organization of bovine DNA. Chromosoma (Berl.) 76:277-294.

McCutchan, Thomas F. and Maxine F. Singer. 1981. DNA sequences similar to those around the simian virus 40 origin of replication are present in the monkey genome. Proc. Natl. Acad. Sci. (USA) 78:95-99.

Meselson, Matthew, Franklin W. Stahl, and Jerome Vinograd. 1957. Equilibrium sedimentation of macromolecules in density gradients. Proc. Natl. Acad. Sci. (USA) 43:581-588.

Messing, Joachim, Roberto Crea, and Peter Seeburg. 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. 9:309-321

Orgel, L. E. and F. H. C. Crick. 1980. Selfish DNA: the ultimate parasite. Nature 284:604-607.

Pardue, Mary Lou and Joseph G. Gall. 1970. Chromosomal localization of mouse satellite DNA. Science 168:1356-1358.

Perlman, S., C. Phillips, and J. O. Bishop. 1976. A study of foldback DNA. Cell 8:33-42.

Reddy, Ramachandra, Dale Henning, and Harris Busch. 1981a. The primary nucleotide sequence of U4 RNA. J. Biol. Chem. 256:3532-3538.

Reddy, Ramachandra, Dale Henning, Paul Epstein, and Harris Busch. 1981b. Primary and secondary structure of U2 snRNA. Nucleic Acids Res. 9:5645-5658.

Rigby, Peter W., Marianne Dieckmann, Carl Rhodes, and Paul Berg. 1977. Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113:237-251.

Rinehart, Frank P., Vincent L. Wilson, and Carl W. Schmid. 1977. Fractionation and the sequence organization of human DNA by preparative $HgCl_2/Cs_2SO_4$ density gradients. Biochim. Biophys. Acta 474:69-81.

Rubin, Carol M., Catherine M. Houck, Prescott L. Deininger, Theodore Friedmann, and Carl W. Schmid. 1980. Partial nucleotide sequence of the 300 nucleotide interspersed repeated human DNA sequences. Nature 284:372-374.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. (USA) 74:5463-5467.

Schmid, Carl W. and Prescott L. Deininger. 1975. Sequence organization of the human genome. Cell 6:345-358.

Schmid, Carl W. and Warren R. Jelinek. 1982. The alu family of dispersed repetitive sequences. Science 216:1065-1070.

Sorber, Herbert A., ed. 1968. Handbook of biochemistry, selected data for molecular biology. Second edition. The Chemical Rubber Company. Cleveland, Ohio. ppH3-H136.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Strobel, Edward, Pamela Dunsmuir, and Gerald M. Rubin. 1979. Polymorphisms in the chromosomal locations of elements of the 412, copia, and 297 dispersed repeated gene families in Drosophila. Cell 17:429-439.

Sutcliffe, J. Gregor. 1978. pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. Nucleic Acids Res. 5:2721-2729.

Taylor, J. H. and S. Watanabe. 1981. In structure and DNA-protein interactions of replication origins. ICN-UCLA Symposia on molecular and cellular biology, eds. D. A. Ray and C. F. Fox. Vol. 22 pp. 597-606. Academic Press, New York.

Ullu, Elisabetta and Marialouisa Melli. 1982. Cloning and characterization of cDNA copies of the 7S RNAs of HeLa cells. Nucleic Acids Res. 10:2209-2223.

Walter, Peter and Gunter Blobel. 1982. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. Nature 299:691-698.

Waring, Michael and Roy J. Britten. 1966. Nucleotide sequence repetition: a rapidly reassociating fraction of mouse DNA. Science 154:791-794.

Watanabe, Yumiko, Toshihiko Tsukada, Mitsue Notake, Shigetada Nakanishi, and Shosaku Numa. 1982. Structural analysis of repetitive DNA sequences in the bovine corticotropin-B-lipotropin precursor gene region. Nucleic Acids Res. 10:1459-1469.

Wu, Ray ed. 1979. Methods in enzymology. Volume 68. Recombinant DNA. Academic Press, New York. 555pp.

ACKNOWLEDGMENTS

The help and support from many people has enabled me to finish the work in this dissertation. I want to thank the following people for all they've done.

My major professor, Dr. John Mayfield, "Thanks" isn't a big enough word.

Susan Barnum, for sanity

Nancy and David Rosen, for taking in a stray

Mr. and Mrs. Frederick Schneider, for being my Ames "parents"

Mr. and Mrs. Neil Richardson, for a calm sanctuary and low grocery bills

Jim Bassuk, Holly Van Oort, Kevin Kraus, Steve Libby, and Tom McMullen, for being co-conspirators in comiseration

Herbert Brothers, for all the gels - "Hey, Herb"

"Uncle" Al, for all the malts I made